UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

FACULTAD DE ESTUDIOS SUPERIORES ACATLÁN

Cálculo diferencial e integral de una y varias variables aplicado a estadística, probabilidad y finanzas

TESINA

QUE PARA OBTENER EL TÍTULO DE

LICENCIADO EN ACTUARÍA

PRESENTA

JIMÉNEZ GUTIÉRREZ GUSTAVO

Asesor: ALEJANDRO MARTÍNEZ IRENEO

Junio 2016

Santa Cruz Acatlán, Naucalpan, Estado de México

Agradecimientos

A mis padres y ejemplos de vida, Sara y Ramón, que con su apoyo, enseñanzas, paciencia y sobre todo cariño me han acompañado en cada paso y proyecto de mi vida, su confianza es invaluable y me motiva a seguirme superando. Agradezco a la vida haberme hecho su hijo.

A Cynthia, mi pequeña hermana que siempre me ha demostrado su grandeza con tanto cariño y me obliga a tratar de ser un buen ejemplo.

A Alicia, mi novia, ya que siempre eh contado con ella y le eh aprendido tanto, ese objetivo que teníamos en 2005 al conocernos lo hemos concretado y juntos, gracias.

A mi asesor, Alejandro Martínez Ireneo, que me ha compartido su conocimiento y enseñanzas y que me acompaño en mi recorrer como estudiante y ahora en este camino para convertirme en profesionista, por creer en mí y permitirme dar clases a su lado.

A mis sinodales que dedicaron tiempo en la revisión de este trabajo y contribuyeron en él.

A cada uno de mis maestros, familiares y amigos que de alguna u otra manera me han acompañado y apoyado teniendo una influencia importante en mi vida.

A la UNAM por brindarme la formación académica que poseo y por ser la institución de la cual me siento orgulloso de pertenecer.

Índice general

	Página
Introducción	III
Capítulo 1. Preliminares	1
1.1. Funciones de distribución	1
1.2. Transformaciones	7
Capítulo 2. Marco teórico de variables aleatorias en el espacio	13
2.1. Distribuciones conjuntas y marginales	13
2.2. Momentos de una variable aleatoria e independencia	16
2.3. Probabilidad condicional y esperanza condicional	18
2.4. Transformaciones multidimensionales	19
2.5. Estadísticos de orden	21
2.6. Estimador de máxima verosimilitud	21
Capítulo 3. Ejercicios teóricos	23
3.1. Distribución conjunta y marginales	23
3.2. Momentos e independencia de variables aleatorias	31
3.3. Probabilidad y esperanza condicional	37
3.4. Transformaciones multidimensionales	40
3.5. Estadísticos de orden	46
3.6. Máxima verosimilitud.	50
Capítulo 4. Ejercicios aplicados	53
4.1. Seguros	53
4.2. Economía	59
4.3. Meteorología	65
4.4. Varios	67
Conclusiones	72
Apéndice A	73
Referencias	76

Introducción

Una de las cualidades de la probabilidad y la estadística es que puede describir patrones y comportamientos sobre algún o algunos sucesos, procesos, actitudes, fenómenos y una gran variedad de eventos, por tal motivo es y ha sido muy importante durante mucho tiempo.

Dichos comportamientos y patrones por si solos, solo son información que no dice nada, es hasta que por medio del uso de técnicas matemáticas, lógicas y a su vez el área de estudio del fenómeno, que se pueden empezar a generar modelos, establecer parámetros, crear estadísticos, inferencias, hipótesis y gran cantidad de datos que tratarán de explicar de la manera más exacta posible nuestro fenómeno.

Mientras mayor sea la muestra analizada, mejores y más certeras aproximaciones se tendrán, se sabe que no serán resultados del todo exactos, pero sí presentarán una buena aproximación a los resultados reales.

Está necesidad de modelar los fenómenos surge del deseo que tiene el ser humano de tratar de comprender con el mayor grado de exactitud, todos los acontecimientos que suceden a su alrededor y por los que resulta afectado, pero eso que no logra entender, es lo que lo intriga y lo que lo lleva a un amplio estudio del tema en cuestión, de esta manera y con base en sus resultados puede entender las razones por las cuales sucede ese fenómeno, para qué pasa, desde cuándo pasa, hasta cuándo pasará y una gran variedad de preguntas alrededor del fenómeno.

Mientras más preguntas se puedan resolver respecto al fenómeno y más datos u observaciones se tengan de éste, se estará comprendiendo de la mejor manera el o los comportamientos que éste sigue o presenta, ese es el objetivo que la probabilidad y estadística han seguido desde siempre y cada vez, éste se está volviendo una necesidad en la vida y entorno del ser humano, una herramienta que se desea dominar.

¿Por qué una necesidad?, la respuesta es simple, tratamos de comprender de forma exacta para poder predecir de manera casi exacta algo que se presentará en el "futuro", si el futuro fuese algo cierto, no tendría sentido el objetivo de la probabilidad y estadística, pero como es algo incierto, término conocido en el campo de dichas materias como "aleatorio", es ahí donde tratamos de predecir o reproducir resultados que esperamos que sucedan.

La toma de decisiones sobre algo incierto es el mayor uso que tienen la probabilidad y estadística, ya que esa decisión tomada hoy, va a tener un impacto en el futuro, el objetivo de esta decisión la mayoría de las veces pretende generar un beneficio para nosotros, para minimizar riesgos, maximizar ganancias, minimizar costos, advertir sobre posibles sucesos, alertar sobre algún problema que se pueda presentar o en su caso, vislumbrar éxito venidero.

La presencia de las matemáticas en la vida del ser humano es esencial y lo es también en el tema de la probabilidad y la estadística, ya que en base a ellas es como se pueden construir las bases que determinaran y darán forma a las herramientas estadísticas y probabilísticas que a su vez, ayudarán a la toma de decisiones, sea cual sea el objetivo que esta decisión busque o tenga por finalidad.

Es por eso que para poder entender a estas dos materias, resulta indispensable conocer y dominar una gran cantidad de técnicas matemáticas y además, saber sus usos y significados para después llevarlos a la creación de modelos probabilísticos y estadísticos que con base en experiencias y datos pasados, tratarán de predecir comportamientos futuros.

El objetivo que este trabajo persigue es mostrar al lector el por qué es tan importante conocer la rama de las matemáticas, para posteriormente comprender el campo de estudio de la probabilidad y la estadística y al final, analizar, modelar y dar soluciones a problemas que estas materias envuelven. Cada capítulo describe su objetivo y plantea el contenido que se manejará para que el lector, conforme avance la lectura, comprenda más la relación que existe entre las matemáticas y dichas materias.

Dado que el campo de trabajo del actuario se encuentra principalmente en el ramo de las finanzas, los seguros y el riesgo, es muy importante que conozca y domine las técnicas que se exponen en esta obra, para que al momento de enfrentarse a diversos problemas, los pueda atacar y resolver de la mejor y más óptima manera y así, obtener los resultados que este persiga. La mayoría de las decisiones están involucradas con finanzas o aspectos que de alguna manera tendrán impacto en las finanzas de una entidad o persona.

CAPÍTULO 1

Preliminares

Con el objetivo de estudiar y comprender eventos de la vida real, la teoría de la probabilidad genera modelos probabilísticos, los cuales se caracterizan por ser aleatorios y se rigen por ciertas propiedades, las cuales son estudiadas para el mejor conocimiento del modelo.

El curso de probabilidad I se relaciona directamente con el cálculo diferencial e integral de una variable debido a que, para entender el comportamiento de las variables aleatorias, se requiere del uso de teoremas, fórmulas, funciones y herramientas que el cálculo proporciona para demostrar y justificar, axiomas y propiedades que las variables aleatorias unidimensionales cumplen.

Las variables aleatorias unidimensionales tienen una estrecha relación con el cálculo de una variable ya que a la variable aleatoria se le modela como una función que puede ser continua o discreta (escalonada). Por ser funciones, estas cumplen con ciertos teoremas y propiedades que el cálculo otorga a ellas.

En este capítulo se analizan técnicas y conceptos ligados entre las variables aleatorias unidimensionales y el cálculo de una variable, así como resultados importantes, que arroja la teoría de probabilidades, que nos ayudaran y serán útiles en el caso multidimensional.

1.1. Funciones de distribución

A cada variable aleatoria X, se le asocia una función llamada función de distribución acumulativa de X

Definición 1.1.1 La función de distribución acumulativa de una variable aleatoria X, denotada por $F_X(x)$, queda definida por $F_X(x) = P_X(X \leq x)$, para toda x.
Una variable aleatoria X es continua si $F_X(x)$ es una función de x continua. Una variable aleatoria X es discreta si $F_X(x)$ es una función de salto de x.

Definición 1.1.2 Funciones masa y densidad de probabilidad
La función masa de probabilidad o fmp, $f_X(x)$, de una variable aleatoria discreta X es dada por:

$$f_X(x) = P(X = x) \;\; para\, toda\, x.$$

$$F_X(x) = P_X(X \leq x) = \sum_{k=1}^{x} f_X(k)$$

La función densidad de probabilidad o fdp, $f_X(x)$, de una variable aleatoria continua X es dada por:

$$f_X(x) = \frac{d}{dx} F_X(x)$$

Teorema 1.1.3[1] Una función $f_X(x)$ es una fdp (o fmp) de una variable aleatoria X sí y soló si

(1) $f_X(x) \geq 0 \quad para\, toda\, x.$

(2) $\sum_x f_X(x) = 1 \;\, (fmp) \quad o \quad \int_{-\infty}^{\infty} f_X(x)dx = 1 \;\, (fdp).$

Ejemplo 1.1.4 Distribución discreta Poisson.

Suponga que se tiene una distribución binomial$(n, \frac{\lambda}{n})$ con fmp de la forma

$$f_X(x) = \binom{n}{x} \left(\frac{\lambda}{n}\right)^x \left(1 - \frac{\lambda}{n}\right)^{n-x} = \frac{n!}{(n-x)!x!} \frac{\lambda^x}{n^x} \left(1 - \frac{\lambda}{n}\right)^{n-x}.$$

Se hallará $\lim_{n\to\infty} f_X(x)$.

$$\begin{aligned}
\lim_{n\to\infty} f_X(x) &= \lim_{n\to\infty} \binom{n}{x} \left(\frac{\lambda}{n}\right)^x \left(1 - \frac{\lambda}{n}\right)^{n-x} \\
&= \lim_{n\to\infty} \frac{n!}{(n-x)!x!} \frac{\lambda^x}{n^x} \left(1 - \frac{\lambda}{n}\right)^{n} \left(1 - \frac{\lambda}{n}\right)^{-x}.
\end{aligned}$$

Al aplicar la propiedad de los límites $\lim_{x\to a} [f(x) \cdot g(x)] = \lim_{x\to a} f(x) \cdot \lim_{x\to a} g(x)$, se tiene

$$\begin{aligned}
\left(1 - \frac{\lambda}{n}\right)^n &= exp\left\{n \cdot log\left(1 - \frac{\lambda}{n}\right)\right\} \\
\lim_{n\to\infty} \left(1 - \frac{\lambda}{n}\right)^n &= \lim_{n\to\infty} exp\left\{n \cdot log\left(1 - \frac{\lambda}{n}\right)\right\} \\
&= exp\left\{\lim_{n\to\infty} n \cdot log\left(1 - \frac{\lambda}{n}\right)\right\}
\end{aligned}$$

Al trabajar el límite por separado se obtiene

$$\lim_{n\to\infty} n \cdot log\left(1 - \frac{\lambda}{n}\right) = \lim_{n\to\infty} \frac{log\left(1 - \frac{\lambda}{n}\right)}{\frac{1}{n}}$$

[1] Casella, G., Berger, R.L. (2002). Statistical inference. Pacific Grove CA: Duxbury. (p. 36)

$(al\ utilizar\ la\ regla\ de\ L'Hôpital)$

$$
\begin{aligned}
&= \lim_{n\to\infty} \frac{\frac{\frac{\lambda}{n^2}}{\left(1-\frac{\lambda}{n}\right)}}{\frac{-1}{n^2}} \\
&= \lim_{n\to\infty} \frac{\frac{\lambda}{n^2}}{\left(1-\frac{\lambda}{n}\right)\left(\frac{-1}{n^2}\right)} \\
&= \lim_{n\to\infty} \frac{-\lambda}{\left(1-\frac{\lambda}{n}\right)} \\
&= -\lambda
\end{aligned}
$$

Por lo tanto

$$\lim_{n\to\infty}\left(1-\frac{\lambda}{n}\right)^n = e^{-\lambda} \tag{1.1.1}$$

Por otra parte

$$\lim_{n\to\infty}\left(1-\frac{\lambda}{n}\right)^{-x} = \lim_{n\to\infty}\frac{1}{\left(1-\frac{\lambda}{n}\right)^x} = \frac{1}{1} = 1 \tag{1.1.2}$$

Se tiene que

$$
\begin{aligned}
\lim_{n\to\infty}\frac{n!}{(n-x)!n^x} &= \lim_{n\to\infty}\frac{1\times\ldots\times(n-x-1)\times(n-x)\times\ldots\times(n-1)\times n}{1\times 2\times 3\times\ldots\times(n-x-1)\times(n-x)}\left(\frac{1}{n^x}\right) \\
&= \lim_{n\to\infty}\frac{(n-x+1)\times\ldots\times(n-1)\times n}{n^x}.
\end{aligned}
$$

Es fácil notar que el numerador tiene x términos y x términos tiene el denominador por lo tanto se puede escribir este límite de la siguiente forma

$$
\begin{aligned}
\lim_{n\to\infty}\frac{n!}{(n-x)!n^x} &= \lim_{n\to\infty}\left(\frac{n}{n}\right)\left(\frac{n-1}{n}\right)\left(\frac{n-2}{n}\right)\cdots\left(\frac{n-x+1}{n}\right) \\
&= 1
\end{aligned} \tag{1.1.3}
$$

Ya que cada uno de estos x términos se aproxima a 1 cuando n tiende a ∞

Por lo tanto se tiene que

$$
\begin{aligned}
\lim_{n\to\infty}\frac{n!}{(n-x)!x!}\frac{\lambda^x}{n^x}\left(1-\frac{\lambda}{n}\right)^{n-x} &= \lim_{n\to\infty}\frac{n!}{(n-x)!n^x}\left(1-\frac{\lambda}{n}\right)^{-x}\frac{\lambda^x}{x!}\left(1-\frac{\lambda}{n}\right)^n \\
&= \frac{\lambda^x}{x!}\lim_{n\to\infty}\frac{n!}{(n-x)!n^x}\cdot\lim_{n\to\infty}\left(1-\frac{\lambda}{n}\right)^{-x}\cdot\lim_{n\to\infty}\left(1-\frac{\lambda}{n}\right)^n
\end{aligned}
$$

$(al\ utilizar\ (1.1.1),\ (1.1.2)\ y\ (1.1.3))$

$$= \frac{\lambda^x}{x!} e^{-\lambda}$$

Que coincide con la fmp de una distribución Poisson(λ).

Una distribución Poisson tiene fmp

$$f_X(x) = \frac{\lambda^x}{x!} e^{-\lambda};\ \ x = 0, 1, 2, \ldots;\ \ 0 \leq \lambda < \infty$$

La distribución Poisson tiene una amplia gama de aplicaciones principalmente en el modelado de diferentes tipos de experimentos. Uno de los más importantes es el modelar el fenómeno que conlleva el tiempo de espera para que ocurra un evento, el número de ocurrencias en un intervalo de tiempo dado puede algunas veces ser modelado por una distribución Poisson.

Se demostrará que $f_X(x) = \frac{\lambda^x}{x!} e^{-\lambda}$ cumple con ser fmp.
La función *exponencial*[2] (e^{α}) es mayor o igual a cero. Por lo tanto

$$e^{-\lambda} \geq 0$$

Se sabe que [3]$\left(\lambda^x = e^{xlog(\lambda)}\right)$, con lo que

$$\lambda^x \geq 0$$

Y por último de la definición $x \varepsilon \{Z^+ \cup 0\}$, de estos 3 hechos se puede garantizar que

$$f_X(x) = \frac{\lambda^x}{x!} e^{-\lambda} \geq 0$$

.
Ahora se demostrará que $\sum\limits_{x=0}^{\infty} f_X(x) = 1$.
Por la expansión de Taylor de (e^x)[4]

$$e^x = \sum_{i=0}^{\infty} \frac{x^i}{i!},$$

por lo tanto

$$\sum_{x=0}^{\infty} \frac{\lambda^x}{x!} e^{-\lambda} = e^{-\lambda} \sum_{x=0}^{\infty} \frac{\lambda^x}{x!} = e^{-\lambda} e^{\lambda} = 1.$$

Queda así demostrado que $f_X(x)$ es fmp.

[2] Spivak, M. (2003). Calculus, 2^a ed. México: Reverté (p.471)

[3] Spivak, M. (2003). Calculus, 2^a ed. México: Reverté (p.474)

[4] Courant, R. Fritz, J. (1999). Introducción al cálculo y al análisis matemático Vol. I. México: Limusa (p. 471)

Ejemplo 1.1.5 Función de distribución continua $Normal(\mu, \sigma^2)$.

La fdp de una variable aleatoria $Normal(\mu, \sigma^2)$, es de la forma

$$f(x;\mu,\sigma^2) = \frac{1}{\sqrt{2\pi}\sigma} e^{\frac{-(x-\mu)^2}{2\sigma^2}}, \ -\infty < x < \infty, \ -\infty < \mu < \infty, \ \sigma > 0$$

Se analizarán las propiedades de la fdp de una variable aleatoria $Normal(\mu, \sigma^2)$.

$$f(x;\mu,\sigma^2) = \frac{1}{\sqrt{2\pi}\sigma} e^{\frac{-(x-\mu)^2}{2\sigma^2}}.$$

$$\frac{df(x;\mu,\sigma^2)}{dx} = \frac{1}{\sqrt{2\pi}\sigma} e^{\frac{-(x-\mu)^2}{2\sigma^2}} \cdot \frac{-2\,(x-\mu)}{2\sigma^2} = \frac{(\mu - x)}{\sqrt{2\pi}\sigma^3} e^{\frac{-(x-\mu)^2}{2\sigma^2}},$$

se igualan a cero y se obtienen las raíces de x

$$\frac{(\mu - x)}{\sqrt{2\pi}\sigma^3} e^{\frac{-(x-\mu)^2}{2\sigma^2}} = 0 \Longrightarrow \frac{\mu}{\sqrt{2\pi}\sigma^3} e^{\frac{-(x-\mu)^2}{2\sigma^2}} = \frac{x}{\sqrt{2\pi}\sigma^3} e^{\frac{-(x-\mu)^2}{2\sigma^2}} \Longrightarrow x_1 = \mu \ y \ x_2 = \infty.$$

Se obtiene la segunda derivada de la fdp y se evalua en x_1

$$\begin{aligned} \frac{d^2 f(x;\mu,\sigma^2)}{d^2 x} &= \frac{(\mu - x)}{\sqrt{2\pi}\sigma^3} e^{\frac{-(x-\mu)^2}{2\sigma^2}} \cdot \frac{-2\,(x-\mu)}{2\sigma^2} + e^{\frac{-(x-\mu)^2}{2\sigma^2}} \left(\frac{-1}{\sqrt{2\pi}\sigma^3} \right) \\ &= e^{\frac{-(x-\mu)^2}{2\sigma^2}} \left[\frac{(\mu - x)^2}{\sqrt{2\pi}\sigma^5} - \frac{1}{\sqrt{2\pi}\sigma^3} \right] \Bigg|_{x_1 = \mu} = \frac{-1}{\sqrt{2\pi}\sigma^3}. \end{aligned}$$

Al igualar la segunda derivada a cero se obtienen los puntos críticos

$$e^{\frac{-(x-\mu)^2}{2\sigma^2}} \left[\frac{(\mu - x)^2}{\sqrt{2\pi}\sigma^5} - \frac{1}{\sqrt{2\pi}\sigma^3} \right] = 0 \Longrightarrow e^{\frac{-(x-\mu)^2}{2\sigma^2}} \frac{(\mu - x)^2}{\sqrt{2\pi}\sigma^5} = e^{\frac{-(x-\mu)^2}{2\sigma^2}} \frac{1}{\sqrt{2\pi}\sigma^3} \Longrightarrow$$

$$\frac{(\mu - x)^2}{\sqrt{2\pi}\sigma^5} = \frac{1}{\sqrt{2\pi}\sigma^3} \Longrightarrow (\mu - x)^2 = \sigma^2$$

$$\therefore x_1 = \mu + \sigma \ \ y \ \ x_2 = \mu - \sigma.$$

Se procede a obtener la tercera derivada de la fdp.

$$\frac{d^3 f(x;\mu,\sigma^2)}{d^3 x} = e^{\frac{-(x-\mu)^2}{2\sigma^2}} \left[\frac{(\mu - x)^2}{\sqrt{2\pi}\sigma^5} - \frac{1}{\sqrt{2\pi}\sigma^3} \right] \cdot \frac{-2\,(x-\mu)}{2\sigma^2} + \cdots$$

$$\cdots + e^{\frac{-(x-\mu)^2}{2\sigma^2}} \cdot \frac{-2(\mu - x)}{\sqrt{2\pi}\sigma^5} = e^{\frac{-(x-\mu)^2}{2\sigma^2}} \left[\frac{(\mu - x)^3}{\sqrt{2\pi}\sigma^7} - \frac{\mu - x}{\sqrt{2\pi}\sigma^5} - \frac{\sqrt{2}(\mu - x)}{\sqrt{\pi}\sigma^5} \right].$$

Al evaluar la tercer derivada en los puntos críticos de la segunda derivada se tiene

$$\left.\frac{d^3 f(x;\mu,\sigma^2)}{d^3x}\right|_{x_1=\mu+\sigma} = e^{\frac{-1}{2}} \left[\frac{1}{\sqrt{2\pi}\sigma^4} + \frac{\sqrt{2}}{\sqrt{\pi}\sigma^4} - \frac{1}{\sqrt{2\pi}\sigma^4} \right] = \frac{\sqrt{2}}{\sqrt{\pi}\sigma^4} e^{\frac{-1}{2}} \neq 0,$$

$$\left.\frac{d^3 f(x;\mu,\sigma^2)}{d^3x}\right|_{x_2=\mu-\sigma} = e^{\frac{-1}{2}} \left[\frac{1}{\sqrt{2\pi}\sigma^4} - \frac{1}{\sqrt{2\pi}\sigma^4} - \frac{\sqrt{2}}{\sqrt{\pi}\sigma^4} \right] = -\frac{\sqrt{2}}{\sqrt{\pi}\sigma^4} e^{\frac{-1}{2}} \neq 0.$$

Por lo tanto se tiene que en $x = \mu$ la fdp tiene su único máximo, esto afirma que la media μ también es la moda de la distribución, en $x_1 = \mu + \sigma$ y $x_2 = \mu - \sigma$ la fdp ubica sus puntos de inflexión y por último se nota que

$$f(\mu + a;\mu,\sigma^2) = \frac{1}{\sqrt{2\pi}\sigma} e^{\frac{-(a)^2}{2\sigma^2}} = \frac{1}{\sqrt{2\pi}\sigma} e^{\frac{-(-a)^2}{2\sigma^2}} = f(\mu - a;\mu,\sigma^2),$$

lo que dice que la fdp de una variable aleatoria $Normal(\mu,\sigma^2)$ es simétrica respecto a su media μ.

Se demostrará que la función $f(x;\mu,\sigma^2) = \frac{1}{\sqrt{2\pi}\sigma} e^{\frac{-(x-\mu)^2}{2\sigma^2}}$ es fdp.

Se sabe que la función e es creciente y su rango es $(0,\infty)$ y como está multiplicada por una constante positiva, se tiene que

$$\frac{1}{\sqrt{2\pi}\sigma} e^{\frac{-(x-\mu)^2}{2\sigma^2}} > 0 \quad \forall x \varepsilon R,$$

ahora se demostrará que $\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{\frac{-(x-\mu)^2}{2\sigma^2}} dx = 1$,

Sea $A = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{\frac{-(x-\mu)^2}{2\sigma^2}} dx$ y A que representa el área bajo la curva, por lo tanto se debe demostrar que $A = 1$.

Al hacer la sustitución $z = \frac{(x-\mu)}{\sigma} \implies \frac{dz}{dx} = \frac{1}{\sigma}$ que conlleva a

$$A = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} dz. \tag{1.1.4}$$

Dado que la fdp es positiva es más fácil mostrar que si $A^2 = 1$ entonces $A = 1$. Se puede escribir

$$
\begin{aligned}
A^2 &= \left(\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} dz.\right)^2 \\
&= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} dz \cdot \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}y^2} dy \\
&= \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-\frac{1}{2}(z^2+y^2)} dz dy
\end{aligned}
$$

Cabe resaltar que las variables de integración son soló variables ficticias, por lo tanto el cambio de nombres es permitido.
Al hacer el cambio de variables a coordenadas polares bajo la sustitución

$$z = rsen\theta \quad y = rcos\theta,$$

entonces $z^2 + y^2 = r^2$, $dzdy = rd\theta dr$ y los límites de integración quedan $0 < r < \infty$, $0 < \theta < 2\pi$, por lo tanto

$$
\begin{aligned}
A^2 &= \frac{1}{2\pi} \int_{0}^{\infty} \int_{0}^{2\pi} re^{-\frac{1}{2}r^2} d\theta dr \\
&= \int_{0}^{\infty} re^{-\frac{1}{2}r^2} dr = 1.
\end{aligned}
$$

Definición 1.1.6 Sea Ω cualquier espació con puntos ω y A cualquier subconjunto de Ω. La función indicadora de A, denotada por $I_A(\cdot)$, es la función con dominio Ω y contradominio igual al conjunto que consiste de los dos números reales 0 y 1 definida por

$$I_A(\omega) = \begin{cases} 1 & si\, \omega \in A \\ 0 & si\ \omega \notin A \end{cases} \tag{1.1.5}$$

La función indicadora será usada para indicar subconjuntos de la linea real, por ejemplo

$$I_{([0,1))}(x) = I_{[0,1)}(x) = \begin{cases} 1 & si 0 \leq x < 1 \\ 0 & c.o.c. \end{cases}.$$

1.2. Transformaciones

Teorema 1.2.1[5] Sea X que tiene función de distribución acumulativa $F_X(x)$, sea $Y = g(X)$, y sean χ y Υ tal que

$$\chi = \{x : f_X(x) > 0\} \ \ y \ \ \Upsilon = \{y : y = g(x) \ para\ algún\ x \varepsilon \chi\}$$

entonces.

[5]Casella, G., Berger, R.L. (2002). Statistical inference. Pacific Grove CA: Duxbury. (p. 51)

(1) Si g es una función creciente de χ, $F_Y(y) = F_X(g^{-1}(y))$ para $y \varepsilon \Upsilon$.
(2) Si g es una función decreciente de χ y X es una variable aleatoria continua, $F_Y(y) = 1 - F_X(g^{-1}(y))$ para $y \varepsilon \Upsilon$.

Ejemplo 1.2.2 Suponga que

$$X \sim f_X(x) = \frac{1}{\beta} e^{\frac{-x}{\beta}} \quad 0 \leq x < \infty \text{ y } \beta > 0,$$

defina la transformación $Y = X^{\frac{1}{\gamma}}$, $\gamma > 0$, encuentre la función de distribución acumulativa de y $F_Y(y)$.

X tiene una distribución $exponencial(\frac{1}{\beta})$, por el primer teorema fundamental del cálculo[6] se tiene que la función de distribución acumulativa es

$$F_X(x) = \int_0^x \frac{1}{\beta} e^{\frac{-t}{\beta}} dt = -e^{\frac{-t}{\beta}} \Big|_0^x = 1 - e^{\frac{-x}{\beta}} \quad 0 \leq x < \infty,$$

al hacer $Y = g(x) = X^{\frac{1}{\alpha}}$. Dado que

$$\frac{d}{dx} g(x) = \frac{d}{dx} x^{\frac{1}{\alpha}} = \frac{1}{\alpha} x^{\frac{1}{\alpha}-1} \geq 0, \quad para \;\; 0 \leq x < \infty,$$

se tiene que $g(x)$ es una función creciente. El rango de X es $[0, \infty)$, $x^{\frac{1}{\alpha}}$ tiene el mismo rango $[0, \infty)$ esto es $\Upsilon = [0, \infty)$.

Para $y > 0$,

$$y = x^{\frac{1}{\alpha}} \Longrightarrow x = y^{\alpha}, \; entonces \; g^{-1}(y) = y^{\alpha}.$$

y para $y > 0$, se tiene que

$$F_Y(y) = F_X(g^{-1}(y)) = 1 - exp\left\{-\frac{1}{\beta} y^{\alpha}\right\}.$$

La distribución de la variable aleatoria Y se conoce como la distribución *Weibull* con parámetros $\alpha > 0$, $\beta > 0$ es decir si $Y \sim Weibull(\alpha, \beta)$ su fdp es

$$\begin{aligned} f_Y(y) &= \frac{d}{dy} F_Y(y) \\ &= \frac{d}{dy}\left(1 - exp\left\{-\frac{1}{\beta} y^{\alpha}\right\}\right) \\ &= \frac{\alpha}{\beta} \cdot y^{\alpha-1} \cdot exp\left\{-\frac{1}{\beta} y^{\alpha}\right\}. \end{aligned}$$

[6] Spivak, M. (2003). Calculus, 2^a ed. México: Reverté (p.398)

La distribución *Weibull* modela como se distribuyen los fallos (en sistemas) cuando la tasa de fallos es proporcional a una potencia del tiempo:
Un valor $\gamma < 1$ indica que la tasa de fallos decrece con el tiempo.
Cuando $\gamma = 1$, la tasa de fallos es constante en el tiempo.
Un valor $\gamma > 1$ indica que la tasa de fallos crece con el tiempo.

Teorema 1.2.3[7] Sea X que tiene fdp $f_X(x)$ y sea $Y = g(X)$, donde g es una función monótona. Sean χ y Υ igualmente definidas como en el teorema (1.2.1). Suponga que $f_X(x)$ es continua en χ y que $g^{-1}(y)$ tiene derivada continua en Υ. Entonces la fdp de Y esta dada por

$$f_Y(y) = \frac{d}{dy}F_Y(y) = \begin{cases} f_X(g^{-1}(y))\frac{d}{dy}g^{-1}(y) & si\, g\, es\, creciente \\ -f_X(g^{-1}(y))\frac{d}{dy}g^{-1}(y) & si\, g\, es\, decreciente. \end{cases}$$

Ejemplo 1.2.4 Suponga que X tiene fdp $exponencial(1)$

$$f_X(x) = \frac{1}{e^x} \quad 0 < x < \infty,$$

encuentre la distribución de la variable aleatoria $Y = \alpha - \gamma log(X)$ para $-\infty < \alpha < \infty$ y $\gamma > 0$.

Sea $Y = g(x)$ se analiza el comportamiento de $g(x)$

$$\frac{d}{dx}g(x) = \frac{d}{dx}[\alpha - \gamma log(X)] = -\frac{\gamma}{x} < 0, \quad 0 < x < \infty,$$

por lo tanto $g(x)$ es decreciente, como la función $log(x)$ es monótona creciente[8], al ser $Y = g(x)$ una transformación monótona, se sabe que $Y = g(x)$ también es monótona. El soporte de X es $\chi = (0,\infty)$ y su rango es $\Upsilon = (0,1)$, la función e^x es continua, por lo tanto $f_X(x)$ es continua.

Como $Y = g(x) = \alpha - \gamma log(X)$, entonces

$$g^{-1}(y) = exp\left\{\frac{\alpha - y}{\gamma}\right\} \;\; y \;\; \frac{d}{dy}g^{-1}(y) = -\frac{1}{\gamma}exp\left\{\frac{\alpha - y}{\gamma}\right\}.$$

[7] Casella, G., Berger, R.L. (2002). Statistical inference. Pacific Grove CA: Duxbury. (p. 51)

[8] Bartle, Robert G., (2010). Introducción al análisis matemático de una variable. 3^a ed. México: Limusa Wiley. p(184)

Para $y\varepsilon(-\infty,\infty)$, se tendrá

$$
\begin{aligned}
f_Y(y) &= -f_X(g^{-1}(y))\frac{d}{dy}g^{-1}(y) \\
&= -f_X\left(exp\left\{\frac{\alpha-y}{\gamma}\right\}\right)\left(-\frac{1}{\gamma}exp\left\{\frac{\alpha-y}{\gamma}\right\}\right) \\
&= \frac{\left(exp\left\{-exp\left\{\frac{\alpha-y}{\gamma}\right\}\right\}\right)\left(exp\left\{\frac{\alpha-y}{\gamma}\right\}\right)}{\gamma}
\end{aligned}
$$

La distribución de la variable Y coincide con la distribución *Gumbel* de parámetros $\alpha\varepsilon R$ y $\gamma>0$, por lo tanto si $Y\sim Gumbel(\alpha,\gamma)$ su fdp es

$$
f_Y(y) = \frac{\left(exp\left\{-exp\left\{\frac{\alpha-y}{\gamma}\right\}\right\}\right)\left(exp\left\{\frac{\alpha-y}{\gamma}\right\}\right)}{\gamma}.
$$

Y su función de distribución acumulativa será

$$
\begin{aligned}
F_Y(y) &= \int_{-\infty}^{y}\frac{\left(exp\left\{-exp\left\{\frac{\alpha-t}{\gamma}\right\}\right\}\right)\left(exp\left\{\frac{\alpha-t}{\gamma}\right\}\right)}{\gamma}dt \\
&= exp\left\{-exp-\left\{\frac{y-\alpha}{\gamma}\right\}\right\} = e^{-e^{-\left\{\frac{y-\alpha}{\gamma}\right\}}}.
\end{aligned}
$$

Teorema 1.2.5[9] Sea X que tiene fdp $f_X(x)$ y sea $Y=g(X)$, y defina el espació muestral χ como en el teorema (1.2.1). Suponga que existe una partición, $A_0, A_1,\dots,A_k$, de χ tal que $P(X\varepsilon A_0)=0$ y $f_X(x)$ es continua en cada A_i. Además, suponga que existen funciones $g_1(x),\dots,g_k(x)$, definidas en $A_1,\dots,A_k$, respectivamente, satisfaciendo

$g(x)=g_i(x)$, para $x\varepsilon A_i$,
$g_i(x)$ es monótona en A_i,

el conjunto $\Upsilon=\{y:\ y=g_i(x)\ para\, algun\, x\varepsilon A_i\}$ es la misma para cada $i=1,\dots,k$, y $g_i^{-1}(y)$ tiene derivada continua en Υ, para cada $i=1,\dots,k$.
Entonces

$$
f_Y(y) = \begin{cases} \sum\limits_{i=1}^{k} f_X(g^{-1}(y))\left|\frac{d}{dy}g^{-1}(y)\right| & y\varepsilon\Upsilon \\ 0 & c.o.c \end{cases}.
$$

Ejemplo 1.2.6 Sea X con distribución $Cauchy(\theta,\sigma)$,

$$
f_X(x,\theta,\sigma) = \frac{1}{\pi\sigma}\frac{1}{1+\left(\frac{x-\theta}{\sigma}\right)^2},\ -\infty<x<\infty,\ -\infty<\theta<\infty,\ \sigma>0.
$$

Considere $Y=X^4-2X^2$. Sea $y=g(x)=x^4-2x^2$, entonces

[9]Casella, G., Berger, R.L. (2002). Statistical inference. Pacific Grove CA: Duxbury. (p. 53)

$$y' = 4x^3 - 4x = 4x(x^2 - 1),$$

al obtener los puntos críticos

$$y' = 0 \Longrightarrow x_1 = 0, x_2 = 1 \, y \, x_3 = -1,$$

$$y'' = 12x^2 - 4$$

$$y''(0) = -4, \;\; y''(1) = 8 \; y \; y''(-1) = 8,$$

por lo tanto en el punto $x_1 = 0$ se tiene un máximo y tanto en $x_2 = 1$ y $x_3 = -1$ se encuentran mínimos, que junto a el criterio de la primera derivada se tiene que $g(x) = x^4 - 2x^2$ es monótona en $(-\infty, -1)$, $(-1, 0)$, $(0, 1)$ y $(1, \infty)$, además que $\Upsilon = (0, \infty)$, entonces se pueden tomar

$$\begin{aligned} A_0 &= \{0\}; \\ A_1 = (-\infty, -1), \quad g_1(x) &= x^4 - 2x^2, \quad g_1^{-1}(y) = -\sqrt{\sqrt{y+1}+1}; \\ A_2 = (-1, 0), \quad g_2(x) &= x^4 - 2x^2, \quad g_2^{-1}(y) = -\sqrt{-\sqrt{y+1}+1}; \\ A_3 = (0, 1), \quad g_3(x) &= x^4 - 2x^2, \quad g_3^{-1}(y) = \sqrt{-\sqrt{y+1}+1}; \\ A_4 = (1, \infty), \quad g_4(x) &= x^4 - 2x^2, \quad g_4^{-1}(y) = \sqrt{\sqrt{y+1}+1}. \end{aligned}$$

Nótese que el teorema no puede ser utilizado ya que
$\Upsilon = \{y : y = g_1(x)\} = \{y : y = g_4(x)\} \neq \Upsilon = \{y : y = g_2(x)\} = \{y : y = g_3(x)\}$,
pero se puede restringir a la función

$$f_X(x, \theta, \sigma) = \frac{1}{\pi\sigma} \frac{1}{1 + \left(\frac{x-\theta}{\sigma}\right)^2}, \;\; -\infty < \theta < \infty, \; \sigma > 0.$$

tal que $X \varepsilon \left(-\sqrt{2}, \sqrt{2}\right). \}$. De está forma $g(x) = x^4 - 2x^2$ es monótona en $\left(-\sqrt{2}, -1\right)$, $(-1, 0)$, $(0, 1)$ y $\left(1, \sqrt{2}\right)$, al tomar

$$\begin{aligned} A_0 &= \{0\}; \\ A_1 = \left(-\sqrt{2}, -1\right), \quad g_1(x) &= x^4 - 2x^2, \quad g_1^{-1}(y) = -\sqrt{\sqrt{y+1}+1}; \\ A_2 = (-1, 0), \quad g_2(x) &= x^4 - 2x^2, \quad g_2^{-1}(y) = -\sqrt{-\sqrt{y+1}+1}; \\ A_3 = (0, 1), \quad g_3(x) &= x^4 - 2x^2, \quad g_3^{-1}(y) = \sqrt{-\sqrt{y+1}+1}; \\ A_4 = \left(1, \sqrt{2}\right), \quad g_4(x) &= x^4 - 2x^2, \quad g_4^{-1}(y) = \sqrt{\sqrt{y+1}+1}. \end{aligned}$$

La fdp de Y es

$$f_Y(y) = \frac{1}{\pi\sigma} \frac{1}{1+\left(\frac{-\sqrt{\sqrt{y+1}+1}-\theta}{\sigma}\right)^2} \left| - \frac{1}{4\sqrt{\sqrt{y+1}+1}\sqrt{y+1}} \right| +$$
$$+ \frac{1}{\pi\sigma} \frac{1}{1+\left(\frac{-\sqrt{-\sqrt{y+1}+1}-\theta}{\sigma}\right)^2} \left| \frac{1}{4\sqrt{\sqrt{y+1}+1}\sqrt{y+1}} \right| +$$
$$+ \frac{1}{\pi\sigma} \frac{1}{1+\left(\frac{\sqrt{-\sqrt{y+1}+1}-\theta}{\sigma}\right)^2} \left| - \frac{1}{4\sqrt{\sqrt{y+1}+1}\sqrt{y+1}} \right| +$$
$$+ \frac{1}{\pi\sigma} \frac{1}{1+\left(\frac{\sqrt{\sqrt{y+1}+1}-\theta}{\sigma}\right)^2} \left| \frac{1}{4\sqrt{\sqrt{y+1}+1}\sqrt{y+1}} \right|$$

$$f_Y(y) = \frac{1}{\pi\sigma} \frac{1}{4\sqrt{\sqrt{y+1}+1}\sqrt{y+1}} \left[\frac{1}{1+\left(\frac{-\sqrt{\sqrt{y+1}+1}-\theta}{\sigma}\right)^2} + \frac{1}{1+\left(\frac{\sqrt{-\sqrt{y+1}+1}-\theta}{\sigma}\right)^2} \right] +$$
$$+ \frac{1}{\pi\sigma} \frac{1}{4\sqrt{\sqrt{y+1}+1}\sqrt{y+1}} \left[\frac{1}{1+\left(\frac{\sqrt{-\sqrt{y+1}+1}-\theta}{\sigma}\right)^2} + \frac{1}{1+\left(\frac{\sqrt{\sqrt{y+1}+1}-\theta}{\sigma}\right)^2} \right]$$

Para $0 < y < 1$.

CAPÍTULO 2

Marco teórico de variables aleatorias en el espacio

En diversas ocasiones, al estudiar un evento se requiere conocer más de una de sus características, por lo que no basta con la información que proporciona una sola variable aleatoria, en este caso, se necesita de un estudio conjunto de sus propiedades.

Al analizar grandes muestras, la información que proporciona una variable aleatoria no es suficiente para describir la muestra, por eso es de suma importancia la información de n-variables aleatorias que harán un análisis más robusto. Para llevar a cabo este tipo de análisis, se requiere de teoremas, funciones y procedimientos que genera el cálculo de varias variables.

Por tal motivo, en este capítulo se estudia la relación que existe entre el cálculo de varias variables y modelos probabilísticos que involucran más de una variable aleatoria, estos son llamados modelos multivariados.

Dichos modelos son descritos por funciones en el espacio R^n y estas funciones se rigen por la teoría del cálculo de varias variables. Algunas técnicas y conceptos descritas en el capítulo I son generalizadas en el caso multidimensional.

2.1. Distribuciones conjuntas y marginales

Definición 2.1.1 Sean $X_1,\ldots,X_n$ n variables aleatorias. La función $F_{X_1,\ldots,X_n} : R^n \to [0,1]$, definida por:

$$F_{X_1,\ldots,X_n}(x_1,\ldots,x_n) = P\left[X_1 \leq x_1,\ldots,X_n \leq x_n\right] \tag{2.1.1}$$

es llamada función de distribución acumulativa (fda) o función de distribución conjunta de $X_1,\ldots,X_n$.

La función $F_{X_1,\ldots,X_n}(x_1,\ldots,x_n)$ cumple con tres condiciones

(i) La función $x \mapsto F_{X_1,\ldots,X_n}(x_1,\ldots,x_{j-1},\ x,\ x_{j+1},\ldots,\ x_n)$ es no decreciente y continua por la derecha.

(ii) $\lim\limits_{x\to\infty} F_{X_1,\ldots,X_n}\left(x_1,\ldots,x_{j-1},\ x,\ x_{j+1},\ldots,\ x_n\right) = F_{X_1,\ldots,X_{j-1},X_{j+1},\ldots,X_n}\left(x_1,\ldots,x_{j-1},\ x_{j+1},\ldots,\ x_n\right)$

(iii) $\lim\limits_{x\to-\infty} F_{X_1,\ldots,X_n}\left(x_1,\ldots,x_{j-1},\ x,\ x_{j+1},\ldots,\ x_n\right) = 0.$

Asociada a un vector aleatorio $(X_1,\ldots,X_n)$ con su respectiva fda $F_{X_1,\ldots,X_n}$ está otra función, llamada ya sea función de densidad de probabilidad conjunta $(fdp\,conjunta)$ o función masa de probabilidad conjunta $(fmp conjunta)$. Los términos $(fdp\,conjunta)$ y $(fmp conjunta)$ refieren, respectivamente, a los casos continuos y discretos respectivamente.

Definición 2.1.2 Se dice que las variables aleatorias $X_1,\ldots,X_n$ forman un vector aleatorio discreto si existe una colección finita o infinita numerable de vectores

$$(x_1^{(1)},\ldots x_n^{(1)}),(x_1^{(2)},\ldots x_n^{(2)}),\ldots$$

tales que:

$P\left[X_1=x_1^{(m)},\ldots,X_n=x_n^{(m)}\right]>0,\ \ para\,cualquier\,m.$
$\sum\limits_m P\left[X_1=x_1^{(m)},\ldots,X_n=x_n^{(m)}\right]=1$

En este caso, la función $f_{X_1,\ldots,X_n}:R^n\mapsto[0,1]$ definida por:

$$f_{X_1,\ldots,X_n}(x_1,\ldots x_n)=P\left[X_1=x_1,\ldots,X_n=x_n\right] \tag{2.1.2}$$

es llamada función masa de probabilidad conjunta $(fmp conjunta)$ de $X_1,\ldots,X_n$.

La propiedad de aditividad numerable implica la relación:

$$F_{X_1,\ldots,X_n}(x_1,\ldots x_n)=\sum_{\{(y_1,\ldots,y_n)|y_1\leq x_1,\ldots,y_n\leq x_n\}}f_{X_1,\ldots,X_n}(y_1,\ldots y_n)$$

si $A\subset R^n$ está relación equivale a

$$P\left[(X_1,\ldots,X_n)\in A\right]=\sum_{\{(x_1,\ldots x_n)\in A\}}f_{X_1,\ldots,X_n}(x_1,\ldots x_n).$$

Sea (X,Y) un vector aleatorio discreto bivariado. Entonces la función $f(x,y)$ de R^2 en R definida por $f(x,y)=P(X=x,\,Y=y)$ es la $fmp conjunta$ de (X,Y). Sí es necesario poner énfasis en el hecho de que f es la $fmp conjunta$ de el vector (X,Y) en lugar de algún otro vector, se usará la notación

$$f_{X,Y}(x,y).$$

La $fmp conjunta$ puede ser usada para calcular la probabilidad de algún evento definido en términos de (X,Y). Sea A algún subconjunto de R^2. Entonces

$$P((X,Y)\in A)=\sum_{(x,y)\in A}f(x,y).$$

Definición 2.1.3 Sea $X_1,\ldots,X_n$ un vector aleatorio con función de distribución conjunta $F_{X_1,\ldots,X_n}(x_1,\ldots x_n)$ y $S\subset R^n$ un conjunto abierto tal que

$P[(X_1,\ldots,X_n)\in S]=1.$
Si:

- $F_{X_1,\ldots,X_n}(x_1,\ldots x_n)$ es continua sobre R^n.

- $\frac{\delta^n F_{X_1,\ldots,X_n}}{\delta x_1,\ldots \delta x_n}$ existe y es continua sobre S.

Entonces, el vector aleatorio $(X_1,\ldots,X_n)$ es absolutamente continuo y la función:

$$f_{X_1,\ldots,X_n}(x_1,\ldots x_n)=\begin{cases}\frac{\delta^n F_{X_1,\ldots,X_n}}{\delta x_1,\ldots \delta x_n}(x_1,\ldots x_n) & si\ (x_1,\ldots x_n)\in S\\ 0 & c.o.c\end{cases} \tag{2.1.3}$$

es una función densidad de probabilidad conjunta $(fdp conjunta)$ de $X_1,\ldots,X_n$, en el caso unidimesional de R^1 se le llama función densidad de probabilidad (fdp).

Una $fdp conjunta$ tiene dos propiedades:

(i) $f_{X_1,\ldots,X_n}(x_1,\ldots x_n)\geq 0.$
(ii) $\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} f_{X_1,\ldots,X_n}(x_1,\ldots x_n)=1.$

Para vectores aleatorios bivariados continuos se tiene la importante relación

$$F(x,y)=\int_{-\infty}^{x}\int_{-\infty}^{y} f(s,t)\,dtds.$$

De el teorema fundamental del cálculo, esto implica que

$$\frac{\delta^2 F(x,y)}{\delta x\delta y}=f(x,y)$$

con continuidad en los puntos de (x,y).

Definición 2.1.4 Sea $X_1,\ldots,X_n$ un vector aleatorio discreto con $fmp conjunta$ $f_{X_1,\ldots,X_n}$, se llaman funciones masa de probabilidad marginales a las funciones

$$f_{X_j}(x_j)=\sum_{x_1,\ldots,x_{j-1},x_{j+1},\ldots x_n} f_{X_1,\ldots,X_n}(x_1,\ldots,x_{j-1},x_{j+1},\ldots x_n). \tag{2.1.4}$$

Si (X,Y) es un vector aleatorio discreto bivariado con $fmp conjunta$ $f_{X,Y}(x,y)$. Entonces las funciones masa de probabilidad marginales de X y Y,
$f_X(x)=P(X=x)$ y $f_Y(y)=P(Y=y)$, están dadas por

$$f_X(x) = \sum_{y\in R} f_{X,Y}(x,y) \quad y \quad f_Y(y) = \sum_{x\in R} f_{X,Y}(x,y).$$

Definición 2.1.5 Sea $X_1,\ldots,X_n$ un vector aleatorio continuo con $fdp conjunta$ $f_{X_1,\ldots,X_n}$, se llaman funciones densidad de probabilidad marginales a las funciones

$$\begin{aligned} f_{X_j}\left(x_j\right) &= \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} f_{X_1,\ldots,X_n} \\ &\left(x_1,\ldots,x_{j-1},x_j,x_{j+1},\ldots x_n\right)dx_1\ldots dx_{j-1}dx_{j+1}\ldots dx_n. \end{aligned} \tag{2.1.5}$$

Las funciones densidad de probabilidad marginales de X y Y también son definidas como en el caso discreto con integrales reemplazando sumas. Las funciones densidad de probabilidad marginales pueden ser usadas para calcular probabilidades o esperanzas que involucran soló a la variable aleatoria X o Y. Específicamente para el caso bivariado, las funciones densidad de probabilidad marginales de X y Y están dadas por

$$f_X(x) = \int_{-\infty}^{\infty} f(x,y)\,dy, \quad -\infty < x < \infty,$$

$$f_Y(y) = \int_{-\infty}^{\infty} f(x,y)\,dx, \quad -\infty < y < \infty.$$

2.2. Momentos de una variable aleatoria e independencia

Definición 2.2.1 Esperanza

Sea $(X_1,\ldots,X_k)$ una variable aleatoria discreta $n-dimensional$ con $fmp\ conjunta$ $f_{X_1,\ldots,X_k}(\cdot,\ldots,\cdot)$. El valor esperado de una función $g(\cdot,\ldots,\cdot)$ de la variable aleatoria $n-dimensional$, denotada por $E\left[g\left(X_1,\ldots,X_k\right)\right]$, es definida por

$$E\left[g\left(X_1,\ldots,X_k\right)\right] = \sum g\left(x_1,\ldots,x_k\right) f_{X_1,\ldots,X_k}(x_1,\ldots,x_k), \tag{2.2.1}$$

En el caso de una variable aleatoria continua $n-dimensional$ con $fdp\ conjunta$ $f_{X_1,\ldots,X_k}(\cdot,\ldots,\cdot)$.

$$E\left[g\left(X_1,\ldots,X_k\right)\right] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} g\left(x_1,\ldots,x_k\right) f_{X_1,\ldots,X_k}(x_1,\ldots,x_k). \tag{2.2.2}$$

En particular, si $g\left(x_1,\ldots,x_k\right) = x_i$, entonces

$$E\left[g\left(X_1,\ldots,X_k\right)\right] = E\left[X_i\right].$$

Si $X_1, \ldots, X_n$ son variables aleatorias que tienen valor esperado, entonces

$$E[X_1 + \cdots + X_n] = E[X_1] + \cdots + E[X_n] \tag{2.2.3}$$

Definición 2.2.2 Varianza
Si $g(x_1, ..., x_k) = (x_i - E[X_i])^2$, entonces

$$E[g(X_1, ..., X_k)] = E\left[(x_i - E[X_i])^2\right] = var[X_i]. \tag{2.2.4}$$

la raíz cuadrada positiva de $Var[X]$ es la desviación estándar de X.

La desviación estándar s de una muestra se calcula mediante la fórmula

$$s = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2} \tag{2.2.5}$$

Algunas veces es más cómodo utilizar una fórmula alternativa para la varianza

$$\begin{aligned} Var[X] &= E[X - E[X]]^2 \\ &= E\left[X^2 - 2XE[X] + (E[X])^2\right] \\ &= E\left[X^2\right] - 2(E[X])^2 + (E[X])^2 \\ &= E\left[X^2\right] - (E[X])^2 \end{aligned}$$

Fundamentado en el hecho de que $E[XE[X]] = E[X]E[X] = (E[X])^2$, dado que $E[X]$ es una constante. Por lo tanto

$$Var[X] = E\left[X^2\right] - (E[X])^2. \tag{2.2.6}$$

Si las variables aleatorias $X_1, \ldots, X_n$ son independientes, entonces

$$Var[X_1 + \cdots + X_n] = Var[X_1] + \cdots + Var[X_n] \tag{2.2.7}$$

Definición 2.2.3 Función Generadora de Momentos
Sea X una variable aleatoria con función de distribución acumulativa $F_X(x)$. La función generadora de momentos o fgm, de X, denotada por $M_X(t)$, es

$$M_X(t) = E\left[e^{tX}\right],$$

siempre que la esperanza exista para t en alguna vecindad de 0. Esto implica que hay un $h > 0$ tal que, para toda t en $-h < t < h$, $E\left[e^{tX}\right]$ existe. Si la esperanza no existe en una vecindad de 0, se dice que la función generadora de momentos no existe.

En términos más generales se escribe la fgm de X como

$$M_X(t) = \int_{-\infty}^{\infty} e^{tX} f_X(x) dx \quad si\, X\, es\, continua, \tag{2.2.8}$$

o

$$M_X(t) = \sum_x e^{tX} P[X = x] \quad si\, X\, es\, discreta. \tag{2.2.9}$$

Si X tiene fgm $M_X(t)$, entonces

$$E[X^n] = M_X^{(n)}(0),$$

donde se define

$$M_X^{(n)}(0) = \frac{d^n}{dt^n} M_X(t)\bigg|_{t=0}.$$

Esto es, el n-esimo momento es igual a la n-esima derivada de $M_X(t)$ evaluada en $t = 0$.

Definición 2.2.4 Sea $(X_1, \ldots, X_n)$ una variable aleatoria n-dimensional con $fmp\, conjunta$ $f_{X_1,\ldots,X_n}(x_1, \ldots, x_n)$. $X_1, \ldots, X_n$ son definidas como independientes sí y soló si

$$f_{X_1,\ldots,X_n}(x_1, \ldots, x_n) = \prod_{i=1}^{n} f_{X_i}(x_i) \tag{2.2.10}$$

para todos los valores $(x_1, \ldots, x_n)$ de $(X_1, \ldots, X_n)$.

Sea $(X_1, \ldots, X_n)$ una variable aleatoria n-dimensional con $fdp\, conjunta$ $f_{X_1,\ldots,X_n}(x_1, \ldots, x_n)$. $X_1, \ldots, X_n$ son definidas como independientes sí y soló si

$$f_{X_1,\ldots,X_n}(x_1, \ldots, x_n) = \prod_{i=1}^{n} f_{X_i}(x_i) \tag{2.2.11}$$

para todos los valores $(x_1, \ldots, x_n)$.

2.3. Probabilidad condicional y esperanza condicional

Definición 2.3.1 Sea (X,Y) un vector aleatorio discreto bivariado con $fmp\ conjunta$ $f(x,y)$ y funciones masa de probabilidad marginales $f_X(x)$ y $f_Y(y)$. Para algún x para el que $P(X = x) = f_X(x) > 0$, la *fmp condicional* de Y dado que $X = x$ es la función de y denotada por $f(y \mid x)$ y definida por

$$f(y \mid x) = P(Y = y \mid X = x) = \frac{f(x,y)}{f_X(x)}. \tag{2.3.1}$$

Para algún y para el que $P(Y=y)=f_Y(y)>0$, la $fmp\ condicional$ de X dado que $Y=y$ es la función de x denotada por $f(x\mid y)$ y definida por

$$f(x\mid y)=P(X=x\mid Y=y)=\frac{f(x,y)}{f_Y(y)}.$$

Definición 2.3.2 Sea (X,Y) un vector aleatorio continuo bivariado con $fdp\ conjunta$ $f(x,y)$ y funciones densidad de probabilidad marginales $f_X(x)$ y $f_Y(y)$. Para algún x para el que $P(X=x)=f_X(x)>0$, la $fdp\ condicional$ de Y dado que $X=x$ es la función de y denotada por $f(y\mid x)$ y definida por

$$f(y\mid x)=P(Y=y\mid X=x)=\frac{f(x,y)}{f_X(x)}. \tag{2.3.2}$$

Para algún y para el que $P(Y=y)=f_Y(y)>0$, la $fdp\ condicional$ de X dado que $Y=y$ es la función de x denotada por $f(x\mid y)$ y definida por

$$f(x\mid y)=P(X=x\mid Y=y)=\frac{f(x,y)}{f_Y(y)}.$$

Si X y Y son conjuntamente continuas, entonces para cualquier conjunto A,

$$P\{X\varepsilon A\mid Y=y\}=\int_A\frac{f(x,y)}{f_Y(y)}\quad y\quad P\{Y\varepsilon A\mid X=x\}=\int_A\frac{f(x,y)}{f_X(x)}.$$

Definición 2.3.3 Si $g(Y)$ es una función de Y, entonces el valor esperado condicional de $g(Y)$ dado $X=x$ es denotado por $E(g(Y)\mid x)$ y es dado por

$$E(g(Y)\mid x)=\begin{cases}\sum\limits_y g(y)f(y\mid x) & caso\ discreto\\ \int_{-\infty}^{\infty}g(y)f(y\mid x)\,dy & caso\ continuo.\end{cases} \tag{2.3.3}$$

La varianza de la distribución de probabilidad descrita por $f(y\mid x)$ es llamada varianza condicional de Y dado $X=x$ y es denotada por $Var(Y\mid x)$, de la definición 2.2.6 se tiene

$$Var(Y\mid x)=E\left(Y^2\mid x\right)-(E(Y\mid x))^2. \tag{2.3.4}$$

2.4. Transformaciones multidimensionales

Definición 2.4.1 Transformaciones discretas bivariadas.

Sea (X,Y) un vector aleatorio bivariado con una distribución de probabilidad conocida. Ahora se considera un nuevo vector aleatorio bivariado (U,V) definido por $U=\Phi_1(X,Y)$ y $V=\Phi_2(X,Y)$. Donde $\Phi_1(x,y)$ y $\Phi_2(x,y)$ son algunas funciones especificas.

Si (X,Y) es un vector aleatorio bivariado discreto, entonces hay soló un conjunto contable de valores para el que la $fmp conjunta$ de (X,Y) es positiva. Se llamará a este conjunto $\mathscr{A}$. Se define el conjunto

$$\mathscr{B}=\{(u,v)\,:\,u=\Phi_1\,(x,y)\;y\,v=\Phi_2\,(x,y)\;para\,alg\acute{u}n\;(x,y)\,\varepsilon\mathscr{A}\}\,.$$

Entonces $\mathscr{B}$ es el conjunto contable de posibles valores para el vector aleatorio discreto (U,V). Y si, para cualquier $(u,v)\,\varepsilon\mathscr{B}$, A_{uv} es definido como

$$A_{uv}=\{(x,y)\;\varepsilon\,\mathscr{A}\,:\,\Phi_1\,(x,y)=u\;y\;\Phi_2\,(x,y)=v\}\,.$$

Entonces las $fmp conjunta$ de (U,V), puede ser calculada desde la $fmp conjunta$ de (X,Y) por:

$$f_{(U,V)}\,(u,v)=P\,(U=u,V=v)=P\,((X,Y)\,\varepsilon A_{uv})=\sum_{(x,y)\varepsilon A_{uv}}f_{X,Y}(x,y). \tag{2.4.1}$$

Definición 2.4.2 Transformaciones continuas muldimensionales[1].

Sea $(X_1,...,X_n)$ un vector aleatorio absolutamente continuo con función de densidad conjunta $f_{X_1,...,X_n}$, $\mathscr{A}\subset\mathbb{R}^n$ un conjunto abierto tal que
$P\,[(X_1,...,X_n)\,\varepsilon\mathscr{A}]=1$ y $\Phi\,:\,\mathbb{R}^n\rightarrow\mathbb{R}^n$ una función tal que:

(i) Φ es inyectiva sobre $\mathscr{A}$.

(ii) Si $\varphi\,:\,\Phi(\mathscr{A})\rightarrow\mathbb{R}^n$ es la inversa de Φ y $\varphi_1,...,\varphi_n$ son los componentes de φ, entonces las derivadas parciales $\frac{\delta\varphi_i}{\delta y_j}$ existen y son continuas sobre $\Phi(\mathscr{A})$.

(iii) Si $J_\varphi\,:\,\Phi(\mathscr{A})\rightarrow\mathbb{R}$ es el Jacobiano de φ, entonces $J_\varphi(y)\neq0$ para cualquier $y\,\varepsilon\,\Phi(\mathscr{A})$.

Entonces el vector aleatorio $(Y_1,...,Y_n)=\Phi\,(X_1,...,X_n)$ es absolutamente continuo y una función de densidad conjunta, $f_{Y_1,...,Y_n}$, de $(Y_1,...,Y_n)$ está dada por:

$$f_{Y_1,...,Y_n}\,(y_1,...,y_n)=\left|J_\varphi(y_1,...,y_n)\right|f_{X_1,...,X_n}\,(\varphi\,(y_1,...,y_n)) \tag{2.4.2}$$

para cualquier vector $(y_1,...,y_n)\,\varepsilon\,\mathbb{R}^n$.

Teorema 2.4.3[2] Formula de convolución

Si X y Y son variables aleatorias continuas independientes con fdp $f_X(x)$ y $f_Y(y)$ respectivamente, entonces la fdp de $Z=X+Y$ con $W=X$ es

[1]Nótese que las tres restricciones hacia f son las mismas al teorema de cambio de variable de cálculo integral de varias variables. Claudio Pita Ruiz. (1995). Cálculo Vectorial. México. Prentice Hall. (p. 591)

[2]Casella, G., Berger, R.L. (2002). Statistical inference. Pacific Grove CA: Duxbury. (p. 215)

$$f_Z(z) = \int_{-\infty}^{\infty} f_X(w) f_Y(z-w) dw. \tag{2.4.3}$$

2.5. Estadísticos de orden

Definición 2.5.1 Sean $X_1, \ldots, X_n$ n variables aleatorias. Las variables aleatorias $X_{(1)}, \ldots, X_{(n)}$, que representan el ordenamiento de $X_1, \ldots, X_n$ en forma creciente, son llamados los estadísticos de orden correspondientes a $X_1, \ldots, X_n$.
Los estadísticos de orden son variables aleatorias que satisfacen $X_{(1)} \leq \ldots \leq X_{(n)}$.

Sean $X_1, \ldots, X_n$ n variables aleatorias absolutamente continuas e independientes, con fdp común f. Entonces una $fdp conjunta$ de los estadísticos de orden, $X_{(1)}, \ldots, X_{(n)}$, correspondientes a $X_1, \ldots, X_n$, está dada por:

$$f_{X_{(1)},\ldots,X_{(n)}}(x_1, \ldots, x_n) = \begin{cases} n! f_X(x_1) \cdots f_X(x_n) & -\infty < x_1 < \ldots < x_n < \infty \\ 0 & c.o.c. \end{cases} \tag{2.5.1}$$

Teorema 2.5.2 [3]Sean $X_{(1)}, \ldots, X_{(n)}$ los estadísticos de orden de una muestra aleatoria, $X_1, \ldots, X_n$, de una población continua con fda $F_X(x)$ y fdp $f_X(x)$. Entonces la fdp $conjunta$ de $X_{(i)}$ y $X_{(j)}$, $1 \leq i < j \leq n$, es:

$$\begin{aligned} f_{X_{(i)},X_{(j)}}(u,v) = {} & \frac{n!}{(i-1)!(j-1-i)!(n-j)!} f_X(u) f_X(v) [F_X(u)]^{i-1} \\ & \times [F_X(v) - F_X(u)]^{j-i-1} [1 - F_X(v)]^{n-j} \quad -\infty < u < v < \infty \end{aligned} \tag{2.5.2}$$

2.6. Estimador de máxima verosimilitud

Definición 2.6.1 Sea $f_X(x \mid \theta_1, \ldots, \theta_k)$ que denota la fdp o fmp conjunta de una muestra $\underline{X} = (X_1, \ldots, X_n)$. Entonces, dado que $X = x$ es observado, la función de θ definida por

$$L(\theta, \underline{X}) = L(\theta_1, \ldots, \theta_k \mid x_1, \ldots, x_n) = \prod_{i=1}^{n} f_X(x_i, \theta_1, \ldots, \theta_k). \tag{2.6.1}$$

es llamada función de verosimilitud. La función de la verosimilitud puede ser complicada, por lo que se acostumbra trabajar con el logaritmo de esta verosimilitud, conocida como la log-verosimilitud

$$logL(\theta, \underline{X}) = \ell(\theta_1, \ldots, \theta_k \mid x_1, \ldots, x_n) = \sum_{i=1}^{n} log f_X(x_i, \theta_1, \ldots, \theta_k).$$

[3]Ross, S. M. (2006). A first course in probability theory. New Jersey: Prentice Hall. (p.273)

Ambas funciones están dadas es función de θ para una muestra fija observada $\underline{X}$.
Después de obtener cualquiera de estas funciones, el procedimiento para encontrar el estimador máximo verosímil $\hat{\theta}_{MV}$ de θ es el estándar de cálculo diferencial para encontrar un máximo de una función. Es decir, hay que resolver

$\frac{\delta Log(\theta,\underline{X})}{\delta\theta} = 0$ y verificar que

$\frac{\delta^2 Log(\theta,\underline{X})}{\delta^2\theta} < 0$ en el o los puntos críticos de la primer derivada parcial.

CAPÍTULO 3

Ejercicios teóricos

Después de conocer las propiedades de las variables aleatorias unidimensionales y multidimensionales, es necesario saber aplicarlas en diversos escenarios. De acuerdo a la necesidad o el interés de estudio, se pueden realizar transformaciones permitidas que el vector admita.

Como se vio en el capítulo 2, los modelos probabilísticos multidimensionales son descritos por funciones como la de distribución acumulativa, las marginales y condicionales, sin embargo, en otras ocasiones el objetivo del estudio involucra la suma, resta o producto de más de dos variables aleatorias, es por ello que en estos casos y en otros más se lleva a cabo una transformación.

Cuando el área de estudio comprende muestras de variables aleatorias, resulta útil y en muchos casos indispensable, conocer su orden, sus valores esperados o sus probabilidades, por tal motivo, en este capítulo se realizan aplicaciones teóricas de los teoremas y definiciones expuestas en el capítulo 2, que permiten ampliar el panorama de diversas aplicaciones en los diferentes casos de combinaciones de variables aleatorias.

3.1. Distribución conjunta y marginales

Ejemplo 3.1.1 (función de distribución acumulativa) Sean X, Y dos variables aleatorias independientes con $fdp\ Uniforme(0,L)$ cada una. Hallar la función de distribución acumulativa de la variable $Z = X - Y$.
Como X y Y son independientes, la $fdp conjunta$ es de la forma

$$f_{X,Y}(x,y) = \frac{1}{L^2} \qquad 0 \leq x, y \leq L.$$

Se verificará que efectivamente es $fdp\ conjunta$, con base en la definición 2.1.3 se debe mostrar que se cumple

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\,dxdy = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{1}{L^2}dxdy = 1$$

Como $Z = X - Y$, existen dos casos, cuando $0 \leq x < y \leq L$ y cuando $0 \leq y < x \leq L$, por lo tanto

$$\begin{aligned}
\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\frac{1}{L^2}dxdy &= \int_0^L\int_0^y\frac{1}{L^2}dxdy+\int_0^L\int_0^x\frac{1}{L^2}dydx \\
&= \int_0^L\frac{x}{L^2}dx\Big|_0^y+\int_0^L\frac{y}{L^2}dy\Big|_0^x \\
&= \int_0^L\frac{y}{L^2}dy+\int_0^L\frac{x}{L^2}dx \\
&= \frac{y^2}{2L^2}\Big|_0^L+\frac{x^2}{2L^2}\Big|_0^L \\
&= \frac{1}{2}+\frac{1}{2} \\
&= 1
\end{aligned}$$

Se requiere la función de distribución acumulativa de Z y por (2.1.1)

$$F_Z(z) = P[Z\leq z] = P[X-Y\leq z] = P[x-y\leq z]$$

De nuevo existen dos casos. Cuando $0\leq x<y\leq L\Rightarrow -L\leq -Z\leq 0$ (figura 3.1.1), los límites de integración son definidos por

$$-L<x-y<-z\Longrightarrow y-L<x<y-z$$

$(como\ 0<x)$

$$\Longrightarrow 0<x<y-z$$

$$-L<x-y<-z\Longrightarrow x+z<y<x+L$$

$(si\ x=0)$

$$\Longrightarrow z<y<L$$

En el caso en que $0\leq y<x\leq L\Rightarrow 0\leq Z\leq L$ (figura 3.1.2), es más fácil trabajar con la fórmula $P[Z\leq z]=\frac{1}{2}-P[Z>z]=\frac{1}{2}-P[x-y>z]$. Por lo tanto los límites de integración serán

$$z<x-y<L\Longrightarrow x-L<y<x-z$$

$(como\ 0<y)$

$$\Longrightarrow 0<y<x-z$$

$$z < x - y < L \Longrightarrow z + y < x < L + y$$

$(si\, y = 0)$

$$\Longrightarrow 0 < x < L$$

Con lo que la función de distribución acumulativa de la variable Z es de la forma

$$F_Z(z) = \begin{cases} \int_z^L \int_0^{y-z} \frac{1}{L^2} dxdy & -L < Z - < 0 \\ \frac{1}{2} - \int_z^L \int_0^{x-z} \frac{1}{L^2} dydx & 0 \leq Z < L \end{cases}.$$

$$F_Z(z) = \begin{cases} \frac{(L-z)^2}{2L^2} & -L < Z - < 0 \\ \frac{2Lz - z^2}{2L^2} & 0 \leq Z < L \end{cases}$$

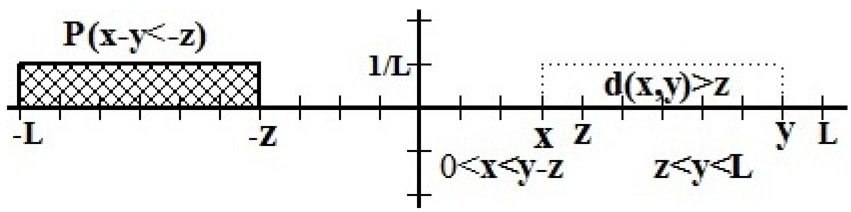

FIGURA 3.1.1. Región del ejemplo 3.1.1 caso $x < y$.

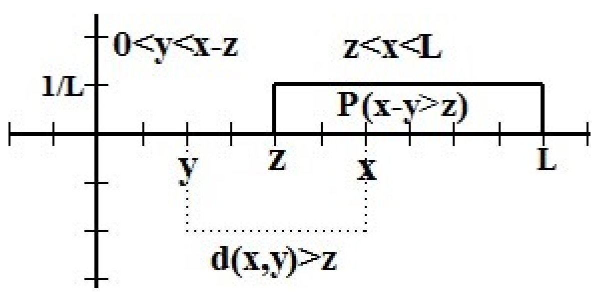

FIGURA 3.1.2. Región del ejemplo 3.1.1 caso $y < x$.

Ejemplo 3.1.2 (función masa de probabilidad conjunta) Sea (X,Y) un vector aleatorio discreto con $fmp\, conjunta$ dada por:

$$f_{X,Y}(x,y) = \begin{cases} kx & si\, x, y \in \{1, \ldots, N\} \\ 0 & c.o.c. \end{cases}.$$

Calcular a) el valor de k para el que $f_{X,Y}(x,y)$ cumple con ser $fmp\ conjunta$, b) $P[X = Y]$ y c) $P[X < Y]$.

a) De la definición 2.1.2 se tiene que

$$\begin{aligned}
\sum_{\{(x,y)|x,y\in\{1,\ldots,N\}\}} f_{X,Y}(x,y) &= Nk \sum_{\{(x,y)|x,y\in\{1,\ldots,N\}\}} x \\
&= Nk \sum_{x=1}^{N} x \\
&= k\frac{N^2(N+1)}{2}. \\
&= 1
\end{aligned}$$

Por lo tanto, $k = \frac{2}{N^2(N+1)}$.

b)

$$\begin{aligned}
P[X=Y] &= \sum_{\{(x,y)|x,y\in\{1,\ldots,N\},x=y\}} f_{X,Y}(x,y) \\
&= k \sum_{\{(x,y)|x,y\in\{1,\ldots,N\},x=y\}} x \\
&= k \sum_{x=1}^{N} x \\
&= \frac{2}{N^2(N+1)} . \frac{N(N+1)}{2} \\
&= \frac{1}{N}.
\end{aligned}$$

c)

$$\begin{aligned}
P[X<Y] &= \sum_{\{(x,y)|x,y\in\{1,\ldots,N\},x<y\}} f_{X,Y}(x,y) \\
&= k \sum_{\{(x,y)|x,y\in\{1,\ldots,N\},x<y\}} x \\
&= k \sum_{x=1}^{N-1} \sum_{y=x+1}^{N} x \\
&= k \sum_{x=1}^{N-1} x(N-x) \\
&= k\left[N\sum_{x=1}^{N-1} x - \sum_{x=1}^{N-1} x^2\right] \\
&= \frac{2}{N^2(N+1)}\left[N\frac{N(N-1)}{2} - \frac{N(N-1)(2N-1)}{6}\right] \\
&= \frac{1}{3}\frac{N-1}{N}
\end{aligned}$$

Ejemplo 3.1.3 (función densidad de probabilidad conjunta) Sean X_1 y X_2 dos variables aleatorias independientes continuas con fdp $exponencial(\lambda)$ cada una. Considérense las transformaciones $U = X_1 + X_2$ y $V = e^{X_1}$. Calcular la $fdp\ conjunta$ de (U,V) y sus $fdp\ marginales$ de U y de V.

Como X_1 y X_2 son variables aleatorias continuas independientes, por (2.2.11), la $fdp conjunta$ de X_1 y X_2 es

$$f_{X_1,X_2}(x_1,x_2) = \lambda^2 e^{-\lambda(x_1+x_2)} \quad 0 \leq x_1,x_2 < \infty$$

Al aplicar (2.4.2), sea el conjunto $\mathscr{A} = \{(x_1,x_2) : 0 < x_1,x_2 < \infty\}$.

Se tiene que $\Phi_1(x_1,x_2) = x_1 + x_2$ y $\Phi_2(x_1,x_2) = e^{x_1}$. Para un valor fijo de x_1

$$\varphi_2 = x_1 = ln(v) \quad 0 < ln(v) < \infty$$

$$\varphi_1 = x_2 = u - ln(v) \quad 0 < u - ln(v) < \infty$$

$$J_{(\varphi_1,\varphi_2)} = \begin{vmatrix} 0 & \frac{1}{v} \\ 1 & -\frac{1}{v} \end{vmatrix} = \left| -\frac{1}{v} \right| = \frac{1}{v}.$$

por lo tanto la $fdp conjunta$ de (U,V) es de la forma

$$f_{U,V}(u,v) = \frac{1}{v} f_{X_1,X_2}(ln(v), u - ln(v)) = \frac{1}{v}\lambda^2 e^{-\lambda u} \quad \begin{matrix} ln(v) < u < \infty \\ 1 < v < e^u \end{matrix} \quad .$$

Para obtener la fdp de U se procede de la siguiente manera

$$f_U(u) = \int_1^{e^u} \frac{1}{v}\lambda^2 e^{-\lambda u} dv = \lambda^2 e^{-\lambda u} dv \int_1^{e^u} \frac{1}{v} dv = \lambda^2 e^{-\lambda u} ln(v) \mid_1^{e^u}$$

$$= \lambda^2 e^{-\lambda u} u \quad 0 < u < \infty.$$

Que es la distribución de una variable aleatoria $Gamma(2,\lambda)$.

Por otra parte la fdp de V es.

$$f_V(v) = \int_{ln(v)}^{\infty} \frac{1}{v}\lambda^2 e^{-\lambda u} du = -\frac{1}{v}\lambda \int_{ln(v)}^{\infty} -\lambda e^{-\lambda u} du = -\frac{1}{v}\lambda e^{-\lambda u} \mid_{ln(v)}^{\infty} =$$

$$= \frac{\lambda}{v^{\lambda+1}} \quad 1 < v < \infty.$$

Ejemplo 3.1.4 (funciones masa de probabilidad marginales) Considérese un vector aleatorio discreto con distribución

$$f_{X,Y,Z}(x,y,z) = c \qquad \begin{array}{c} x = 1, 2, ..., y \\ y = 1, 2, ..., z \\ z = 1, 2, ..., 10 \end{array}.$$

Encontrar las funciones masa de probabilidad marginales $f_X(x)$ y $f_Y(y)$.

Es necesario obtener el valor de c para el que $f_{X,Y,Z}(x,y,z)$ es $fmp conjunta$. Por lo tanto se debe resolver la ecuación para c:

$$\begin{aligned} 1 =& \sum_{z=1}^{10}\sum_{y=1}^{z}\sum_{x=1}^{y} c \\ &= c\sum_{z=1}^{10}\sum_{y=1}^{z} y \\ &= c\sum_{z=1}^{10}\frac{z^2+z}{2} \\ &= \frac{c}{2}\left[\frac{2(10)^3+3(10)^2+10}{6}+\frac{(10)^2+10}{2}\right] \\ &= 220c \end{aligned}$$

$$\therefore c = \frac{1}{220}.$$

Teniendo ya definida la $fmp\ conjunta$.

$$f_{X,Y,Z}(x,y,z) = \frac{1}{220} \qquad \begin{array}{c} x = 1, 2, ..., y \\ y = 1, 2, ..., z \\ z = 1, 2, ..., 10 \end{array}$$

$$\begin{aligned} f_X(x) =& \sum_{z=x}^{10}\sum_{y=x}^{z}\frac{1}{220} \\ &= \frac{1}{220}\sum_{z=x}^{10}(z-x+1) \\ &= \frac{1}{220}\left[\frac{(11-x)(x+10)}{2}\right] - \frac{x}{220}(11-x) + \frac{1}{220}(11-x) \\ &= \left(\frac{(11-x)}{220}\right)\left(\frac{x+10}{2}+x+1\right) \\ &= \frac{(11-x)(12-x)}{440} \quad x = 1, 2, ..., 10 \end{aligned}$$

$$
\begin{aligned}
f_Y(y) &= \sum_{z=y}^{10}\sum_{x=1}^{y}\frac{1}{220} \\
&= \frac{y}{220}\sum_{z=y}^{10} \\
&= \frac{y(11-y)}{220} \qquad y = 1, 2, ..., 10.
\end{aligned}
$$

Ejemplo 3.1.5 (función densidad de probabilidad marginales) Sea

$$
f_{X,Y}(x, y) = \frac{1}{8}\left(y^2 - x^2\right)e^{-y} \qquad \begin{cases} -y \leq x \leq y \\ 0 \leq y < \infty \end{cases}
$$

la $f dp\ conjunta$ de una variable aleatoria (X, Y), hallar las funciones de densidad marginales de X y Y.

Para obtener la marginal de X se integra respecto a y y se observa que
como $-y \leq x \leq y \Rightarrow x \leq |y| \Rightarrow |x| \leq |y| \Rightarrow |y| \geq |x|$ y al ser y positiva, $y \geq |x|$, con lo que

$$
\begin{aligned}
f_X(x) &= \int_{|x|}^{\infty}\frac{1}{8}\left(y^2 - x^2\right)e^{-y}dy \\
&= -\frac{1}{8}\int_{|x|}^{\infty} -y^2e^{-y}dy + \frac{1}{8}x^2\int_{|x|}^{\infty} -e^{-y}dy
\end{aligned}
$$

Sea $u = y^2$, $dv = -e^{-y}dy \Rightarrow du = 2ydy$, $v = e^{-y}$, al integrar por partes se obtiene

$$
\begin{aligned}
f_X(x) &= -\tfrac{1}{8}\left[y^2e^{-y}dy\Big|_{|x|}^{\infty} + \textstyle\int_{|x|}^{\infty} -2ye^{-y}dy\right] + \tfrac{1}{8}x^2e^{-y}dy\Big|_{|x|}^{\infty} \\
&= -\frac{1}{8}\left[-|x|^2e^{|x|} + \int_{|x|}^{\infty} -2ye^{-y}dy\right] - \frac{1}{8}x^2e^{|x|}
\end{aligned}
$$

Al aplicar de nuevo integración por partes con $u = 2y$, $dv = -e^{-y}dy \Rightarrow du = 2ydy$, $v = e^{-y}$

$$
\begin{aligned}
f_X(x) &= -\frac{1}{8}\left[-|x|^2e^{|x|} + 2ye^{-y}\Big|_{|x|}^{\infty} + 2\int_{|x|}^{\infty} -e^{-y}dy\right] - \frac{1}{8}x^2e^{|x|} \\
&= -\frac{1}{8}\left[-|x|^2e^{|x|} - 2|x|e^{-|x|} + 2e^{-y}\Big|_{|x|}^{\infty}\right] - \frac{1}{8}x^2e^{|x|} \\
&= -\frac{1}{8}\left[-|x|^2e^{|x|} - 2|x|e^{-|x|} - 2e^{-|x|}\right] - \frac{1}{8}x^2e^{|x|}
\end{aligned}
$$

$\left(como\ |x|^2 = x^2\right)$

$$= \frac{1}{4}|x|\,e^{-|x|} + \frac{1}{4}e^{-|x|}$$

Se procede a calcular $f_Y(y)$ y está será bajo los límites de integración de x por lo tanto $-y \leq x \leq y$ y entonces

$$\begin{aligned} f_Y(y) &= \int_{-y}^{y} \frac{1}{8}\left(y^2 - x^2\right)e^{-y}dx \\ &= \frac{1}{8}\int_{-y}^{y}\left(y^2 e^{-y} - x^2 e^{-y}\right)dx \\ &= \frac{1}{8}\left[\left(xy^2 e^{-y} - \frac{x^3 e^{-y}}{3}\right)\Big|_{-y}^{y}\right] \\ &= \frac{1}{6}y^3 e^{-y} \end{aligned}$$

Ejemplo 3.1.6 (función marginal mixta) Hallar la función densidad de probabilidad marginal la variable aleatoria Y dado la $fmp\ conjunta$

$$f_{Y,\lambda}(y,\lambda) = \frac{\beta^{\alpha}\lambda^{\alpha-1}e^{-\beta\lambda}e^{-\lambda}\lambda^{y}}{\Gamma(\alpha)\,y!} \qquad \begin{cases} 0 \leq \lambda < \infty \\ y = 0,\ 1,\ 2,\ldots \\ \alpha,\beta > 0 \end{cases}$$

Se procede como indica (2.1.3)

$$\begin{aligned} f_Y(y) &= \int_0^{\infty} \frac{\beta^{\alpha}\lambda^{\alpha-1}e^{-\beta\lambda}e^{-\lambda}\lambda^{y}}{\Gamma(\alpha)\,y!}d\lambda \\ &= \frac{\beta^{\alpha}}{\Gamma(\alpha)\,y!}\int_0^{\infty}\lambda^{y+\alpha-1}e^{-\lambda(\beta+1)}d\lambda \end{aligned}$$

De la relación $\int_0^{\infty} t^{r+k-1}e^{-tp} = p^{r+k}\Gamma(r+k)$[1]

$$f_Y(y) = \frac{\beta^{\alpha}\Gamma(y+\alpha)}{\Gamma(\alpha)\,y!\,(\beta+1)^{y+\alpha}}$$

Como $y\,\varepsilon\,N$, se supondrá que $\alpha\,\varepsilon\,N$ por lo tanto $\alpha + y\,\varepsilon\,N$.

$$f_Y(y) = \frac{\beta^{\alpha}\Gamma(y+\alpha)}{\Gamma(\alpha)\,y!\,(\beta+1)^{y+\alpha}} \quad \alpha = 0,\ 1,\ 2,\ldots$$

Ademas se sabe que $\Gamma(n) = (n-1)!$, entonces

[1] N. N. Lebedev. (1965). Special functions and their applications. Prentice Hall. (p. 27)

$$f_Y(y) = \frac{\beta^{\alpha}(y+\alpha-1)!}{(\alpha-1)\,y!\,(\beta+1)^{y+\alpha}} \quad \alpha = 0,\ 1,\ 2,\ldots$$

Al reordenar términos y del hecho que $\binom{n}{k} = \frac{n!}{k!(n-k)!}$

$$f_Y(y) = \binom{y+\alpha-1}{y}\left(\frac{\beta}{\beta+1}\right)^{\alpha}\left(\frac{1}{\beta+1}\right)^{y} \begin{cases} \alpha = 0,\ 1,\ 2,\ldots \\ y = 0,\ 1,\ 2,\ldots \\ \beta > 0 \end{cases}$$

Entonces la variable aleatoria Y se distribuye como una variable *Binomial Negativa* con parámetros $p = y,\ \alpha,\ \frac{\beta}{\beta+1}$.

La distribución Binomial Negativa puede ser vista como una distribución Poisson, donde el parámetro Poisson es por si mismo una variable aleatoria que se distribuye de acuerdo a una distribución Gamma. Por lo tanto la distribución Binomial Negativa también es conocida como una mezcla Poisson-Gamma.

3.2. Momentos e independencia de variables aleatorias

Ejemplo 3.2.1 (Esperanza, varianza, Fgm de una variable aleatoria discreta) Para la distribución Poisson calcular $E[x]$, $Var[x]$ y $M_X(t)$.
Se tiene que

$$\begin{aligned} E[x] &= \sum_{x=0}^{\infty} x\frac{\lambda^x}{x!}e^{-\lambda} \\ &= \sum_{x=1}^{\infty} x\frac{\lambda^x}{x!}e^{-\lambda} \\ &= \lambda e^{-\lambda}\sum_{x=1}^{\infty}\frac{\lambda^{x-1}}{(x-1)!} \\ &= \lambda e^{-\lambda}e^{\lambda} \\ &= \lambda. \end{aligned}$$

Para la varianza se usará la fórmula $Var[x] = E\left[x^2\right] - E[x]^2$.

$$
\begin{aligned}
E\left[x^2\right] &= \sum_{x=0}^{\infty} x^2 \frac{\lambda^x}{x!} e^{-\lambda} \\
&= \sum_{x=1}^{\infty} x^2 \frac{\lambda^x}{x!} e^{-\lambda} \\
&= \lambda e^{-\lambda} \sum_{x=1}^{\infty} x \frac{\lambda^{x-1}}{(x-1)!} \\
&= \lambda e^{-\lambda} \left(\sum_{x=1}^{\infty} (x-1) \frac{\lambda^{x-1}}{(x-1)!} + \sum_{x=1}^{\infty} \frac{\lambda^{x-1}}{(x-1)!} \right) \\
&= \lambda e^{-\lambda} \left(\lambda \sum_{x=2}^{\infty} \frac{\lambda^{x-2}}{(x-2)!} + \sum_{x=1}^{\infty} \frac{\lambda^{x-1}}{(x-1)!} \right)
\end{aligned}
$$

$(al\ hacer\ m = x-2\ y\ n = x-1)$

$$
\begin{aligned}
&= \lambda e^{-\lambda} \left(\lambda \sum_{m=0}^{\infty} \frac{\lambda^m}{m!} + \sum_{n=0}^{\infty} \frac{\lambda^n}{n!} \right) \\
&= \lambda e^{-\lambda} \left(\lambda e^{\lambda} + e^{\lambda} \right) \\
&= \lambda^2 + \lambda.
\end{aligned}
$$

Por lo tanto

$$Var[x] = \lambda^2 + \lambda - \lambda^2 = \lambda.$$

Por último se tiene que $M_X(t) = E\left[e^{tx}\right]$, con lo que

$$
\begin{aligned}
M_X(t) &= \sum_{x=0}^{\infty} e^{tx} \frac{\lambda^x}{x!} e^{-\lambda} \\
&= e^{-\lambda} \sum_{x=0}^{\infty} \frac{(\lambda e^t)^x}{x!} \\
&= e^{\lambda(e^t-1)}.
\end{aligned}
$$

Nótese el hecho de que:

$$\left.\frac{d}{dt} M_X(t)\right|_{t=0} = \left. e^{\lambda(e^t-1)} \lambda e^{\lambda t} \right|_{t=0} = \lambda = E[x]$$

$$\begin{aligned}
\left.\frac{d^2}{dt^2}M_X(t)\right|_{t=0} &= \left(\lambda e^{\lambda e^t-\lambda+t}\right)\left(\lambda e^t+1\right) \\
&= \left.\lambda^2 e^{\lambda e^t-\lambda+2t}+\lambda e^{\lambda e^t-\lambda+t}\right|_{t=0} \\
&= \lambda^2-\lambda \\
&= E\left[x^2\right].
\end{aligned}$$

Ejemplo 3.2.2 (Esperanza, varianza, Fgm de una variable aleatoria continua) Si X es una variable aleatoria Normal, hallar $M_X(t)$, $E[X]$ y $Var[X]$.

$$\begin{aligned}
M_X(t) &= E\left[e^{tx}\right] \\
&= \frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty} e^{tx}e^{\frac{-(x-\mu)^2}{2\sigma^2}}dx
\end{aligned}$$

$(al\ multiplicar\ por\ e^{\mu t}e^{-\mu t}=1)$

$$\begin{aligned}
&= e^{t\mu}E\left[e^{t(x-\mu)}\right] \\
&= e^{t\mu}\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}\sigma}e^{t(x-\mu)}e^{\frac{-(x-\mu)^2}{2\sigma^2}}dx \\
&= e^{t\mu}\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty} e^{\frac{-1}{2\sigma^2}\left[(x-\mu)^2-2\sigma^2t(x-\mu)\right]}dx \\
&= e^{t\mu}\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty} e^{\frac{-1}{2\sigma^2}\left[(x-\mu)^2-2\sigma^2t(x-\mu)+\sigma^4t^2-\sigma^4t^2\right]}dx \\
&= e^{t\mu}\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty} e^{\frac{-1}{2\sigma^2}\left[\left((x-\mu)-\sigma^2t\right)^2-\sigma^4t^2\right]}dx \\
&= e^{t\mu}e^{\frac{\sigma^2t^2}{2}}\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty} e^{\frac{-1}{2\sigma^2}\left((x-\mu)-\sigma^2t\right)^2}dx,
\end{aligned}$$

Se observa que $\frac{1}{\sqrt{2\pi}\sigma}e^{\frac{-1}{2\sigma^2}\left((x-\mu)-\sigma^2t\right)^2}$ corresponde a la fdp de una variable aleatoria $Normal(\mu+\sigma^2t,\sigma^2)$ por lo tanto

$$\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty} e^{\frac{-1}{2\sigma^2}\left((x-\mu)-\sigma^2t\right)^2}dx = 1,$$

con lo que

$$M_X(t) = e^{t\mu}e^{\frac{\sigma^2t^2}{2}} = e^{t\mu+(\sigma^2t^2)/2}.$$

A partir de que se ah obtenido la fgm se puede calcular tanto la esperanza como la varianza, se tendrá que

$$\begin{aligned} E[X] &= \left.\frac{d}{dt}M_X(t)\right|_{t=0} \\ &= e^{t\mu+(\sigma^2t^2)/2}\left(\mu+\sigma^2 t\right)|_{t=0} \\ &= e^0(\mu+0) \\ &= \mu \end{aligned}$$

Al usar la fórmula de la varianza $Var[X]=E\left[X^2\right]-(E[X])^2$, se procede a encontrar

$$\begin{aligned} E\left[X^2\right] &= \left.\frac{d^2}{d^2t}M_X(t)\right|_{t=0} \\ &= \left.\left(e^{t\mu+(\sigma^2t^2)/2}\right)\left(\sigma^2\right)+e^{t\mu+(\sigma^2t^2)/2}\left(\mu+\sigma^2t\right)^2\right|_{t=0} \\ &= e^0\left(\sigma^2\right)+e^0(\mu)^2 \\ &= \sigma^2+\mu^2. \end{aligned}$$

Con lo que

$$Var[X]=\sigma^2+\mu^2-\mu^2=\sigma^2.$$

Ejemplo 3.2.3 (Variables aleatoria independientes discretas) Sea

$$f_{X,Y}(x,y) \quad = \quad p^{r+s}(1-p)^x\binom{y+r-1}{y}\binom{x-y+s-1}{x-y} \qquad \begin{cases} y<x \\ x-y=1,2,\ldots \\ y=0,1,2,\ldots \end{cases}$$

la $fmp\ conjunta$ de las variables aleatorias X y Y, determinar si X y Y son independientes.

Es necesario encontrar la fmp $marginal$ de X

$$f_X(x) \quad = \quad \sum_{y=0}^{x} p^{r+s}(1-p)^x\binom{y+r-1}{y}\binom{x-y+s-1}{x-y} \quad x=1,2,\ldots$$

Al partir del hecho de que $\left(\binom{k+m-1}{k}=(-1)^k\binom{-m}{k}\right)$[2]

$$\begin{aligned} f_X(x) &= p^{r+s}(1-p)^x\sum_{y=0}^{x}(-1)^y\binom{-r}{y}(-1)^{x-y}\binom{-s}{x-y} \quad x=1,2,\ldots \\ &= p^{r+s}(1-p)^x(-1)^x\sum_{y=0}^{x}\binom{-r}{y}\binom{-s}{x-y} \quad x=1,2,\ldots \end{aligned}$$

[2]Casella, G., Berger, R.L. (2002). Statistical inference. Pacific Grove CA: Duxbury. (p. 95)

Se sabe que $\left(\sum_{w=0}^{k}\binom{n}{w}\binom{m}{k-w}=\binom{n+m}{k}\right)^{3}$

$$\begin{aligned} f_X(x) &= p^{r+s}(1-p)^x(-1)^x\binom{-(r+s)}{x} \quad x=1,2,\ldots \\ &= p^{r+s}(1-p)^x\binom{x+r+s-1}{x} \quad x=1,2,\ldots \end{aligned}$$

Se observa que la variable aleatoria X se distribuye como una variable Binomial negativa con parámetro $(r+s)$.
Ahora se calcula la *fmp marginal* de Y

$$\begin{aligned} f_Y(y) &= \sum_{x-y=1}^{\infty} p^{r+s}(1-p)^x\binom{y+r-1}{y}\binom{x-y+s-1}{x-y} \quad y=0,1,2,\ldots \\ &= p^r(1-p)^y\binom{y+r-1}{y}\sum_{x-y=1}^{\infty} p^s(1-p)^{x-y}\binom{x-y+s-1}{x-y} \quad y=0,1,2,\ldots \end{aligned}$$

$p^s(1-p)^{x-y}\binom{x-y+s-1}{x-y}$ es la *fmp* de una variable $X-Y$ con distribución Binomial Negativa con parámetro s, y como la suma corre por todo su soporte

$\sum_{x-y=1}^{\infty} p^s(1-p)^{x-y}\binom{x-y+s-1}{x-y}=1$

$$f_Y(y)=p^r(1-p)^y\binom{y+r-1}{y} \quad y=0,1,2,\ldots$$

La variable Y también se distribuye como una variable Binomial Negativa con parámetro r.
Finalmente

$$\begin{aligned} f_Y(y)f_X(x) &= p^{2r+s}(1-p)^{x+y}\binom{x+r+s-1}{x}\binom{y+r-1}{y} \\ f_{X,Y}(x,y) &= p^{r+s}(1-p)^x\binom{y+r-1}{y}\binom{x-y+s-1}{x-y} \\ f_Y(y)f_X(x) &\neq f_{X,Y}(x,y) \end{aligned}$$

Y por (2.2.11) las variables aleatorias X y Y no son independientes.
En una secuencia de ensayos los cuales soló pueden tomar el valor de "éxito" con probabilidad p o "fracaso" con probabilidad (1-p), la distribución Binomial Negativa es definida en términos de la variable aleatoria X=número de "fracasos" antes del r-esimo "éxito"

Ejemplo 3.2.4 (Variables aleatorias independientes discretas) Sean X,Y,Z variables aleatorias independientes con distribución $Uniforme(0,L)$. Calcular $P\left[X+Y>\frac{3}{2}Z\right]$.

[3] Harris, J.;Hirst, J.L.; Mossinghoff, M. (2008). Combinatorics and graph theory. Springer. (p. 142)

Como X,Y,Z son independientes, la $fdp conjunta$ de la variable aleatoria (X,Y,Z) es de la forma

$$f_{X,Y,Z}(x,y,z) \quad = \quad \frac{1}{L^3} \quad 0 \leq x,y,z \leq L$$

En la búsqueda de los límites de integración, se tiene la condición inicial
$x+y > \frac{3}{2}z$, al fijar la variable y se obtiene $y > \frac{3}{2}z - x$,
Si $0 < \frac{3}{2}z - x$

$$\begin{aligned}
\frac{3}{2}z - x &< y < L \\
0 < \frac{3}{2}z - x &\Rightarrow 0 < x < \frac{3}{2}z \\
0 < \frac{3}{2}z < L &\Rightarrow 0 < z < \frac{2}{3}L
\end{aligned}$$

Si $\frac{3}{2}z - x < 0$

$$\begin{aligned}
0 &< y < L \\
\frac{3}{2}z - x < 0 &\Rightarrow \frac{3}{2}z < x < L \\
0 < \frac{3}{2}z < L &\Rightarrow 0 < z < \frac{2}{3}L
\end{aligned}$$

Por último si $\frac{3}{2}z - x < L$

$$\begin{aligned}
\frac{3}{2}z - x &< y < L \\
\frac{3}{2}z - x < L &\Rightarrow \frac{3}{2}z - L < x < L \\
0 < \frac{3}{2}z - L &\Rightarrow \frac{2}{3}L < z < L
\end{aligned}$$

Por lo tanto

$$P\left[X+Y>\frac{3}{2}Z\right]=P\left[x+y>\frac{3}{2}z\right]=$$
$$\int_0^{\frac{2}{3}L}\int_0^{\frac{3}{2}z}\int_{\frac{3}{2}z-x}^{L}\frac{1}{L^3}dydxdz+\int_0^{\frac{2}{3}L}\int_{\frac{3}{2}z}^{L}\int_0^{L}\frac{1}{L^3}dydxdz+\int_{\frac{2}{3}L}^{L}\int_{\frac{3}{2}z-L}^{L}\int_{\frac{3}{2}z-x}^{L}\frac{1}{L^3}dydxdz=$$
$$\int_0^{\frac{2}{3}L}\left(\int_0^{\frac{3}{2}z}\int_{\frac{3}{2}z-x}^{L}\frac{1}{L^3}dydx+\int_{\frac{3}{2}z}^{L}\int_0^{L}\frac{1}{L^3}dydx\right)dz+\int_{\frac{2}{3}L}^{L}\int_{\frac{3}{2}z-L}^{L}\int_{\frac{3}{2}z-x}^{L}\frac{1}{L^3}dydxdz=$$
$$\frac{6L^2-1}{9L^3}+\frac{7}{72}.$$

3.3. Probabilidad y esperanza condicional

Ejemplo 3.3.1 (función masa de probabilidad condicional y esperanza condicional discreta)
Sea (X,Y) un vector aleatorio discreto con $fmp\ conjunta$

$$f_{X,Y}(x,y)=\begin{cases}\frac{6}{N\left(N^2-1\right)}(y-x) & \begin{array}{c}x<y\\ \{x,y\}\,\varepsilon\,\{1,\ 2,\ldots,N\}\end{array}\\ 0 & c.o.c\end{cases}$$

Calcular a) $f(x\mid y)$ y $f(y\mid x)$, b) $E(x\mid y)$ y $E(y\mid x)$.
a) El primer paso es encontrar las fmp marginales $f_X(x)$ y $f_Y(y)$.

$$\begin{aligned}f_X(x)&=\sum_{y=x+1}^{N}\frac{6}{N(N^2-1)}(y-x)\\&=\frac{6}{N(N^2-1)}\sum_{y=x+1}^{N}(y-x)\\&=\frac{6}{N(N^2-1)}\left(\frac{n(n+1)}{2}-\frac{x(x+1)}{2}-(n-x)x\right)\\&=\frac{6}{N(N^2-1)}\left(\frac{n^2+n+x^2-x-2nx}{2}\right).\end{aligned}$$

$$\begin{aligned}f_Y(y)&=\sum_{x=1}^{y-1}\frac{6}{N(N^2-1)}(y-x)\\&=\frac{6}{N(N^2-1)}\sum_{x=1}^{y-1}(y-x)\\&=\frac{6}{N(N^2-1)}\left(\frac{y(y-1)}{2}\right).\end{aligned}$$

Con las fmp marginales se obtienen las funciones de distribución condicional

$$
\begin{aligned}
f(x \mid y) &= \frac{f(x,y)}{f_Y(y)} \\
&= \frac{\frac{6}{N\left(N^2-1\right)}(y-x)}{\frac{6}{N\left(N^2-1\right)}\left(\frac{y(y-1)}{2}\right)} \\
&= \frac{2(y-x)}{y(y-1)} \quad \begin{matrix} x<y \\ \{x,\, y\}\,\varepsilon\,\{1,\ 2,\, \ldots,\, N\} \end{matrix}
\end{aligned}
$$

$$
\begin{aligned}
f(y \mid x) &= \frac{f(x,y)}{f_X(x)} \\
&= \frac{\frac{6}{N\left(N^2-1\right)}(y-x)}{\frac{6}{N\left(N^2-1\right)}\left(\frac{n^2+n+x^2-x-2nx}{2}\right)} \\
&= \frac{2(y-x)}{(n-x)(n-x+1)} \quad \begin{matrix} x<y \\ \{x,\, y\}\,\varepsilon\,\{1,\ 2,\, \ldots,\, N\} \end{matrix}
\end{aligned}
$$

b) Dado que $E(x \mid y) = \sum\limits_x x f(x \mid y)$ y $E(y \mid x) = \sum\limits_y y f(y \mid x)$ se tiene que

$$
\begin{aligned}
E(x \mid y) &= \sum_{x=1}^{y-1} \frac{2x(y-x)}{y(y-1)} \\
&= \frac{2}{y(y-1)} \sum_{x=1}^{y-1} xy - x^2 \\
&= \frac{2}{y(y-1)} \left(\frac{y^2(y-1)}{2} - \frac{(y-1)^3}{3} - \frac{(y-1)^2}{2} - \frac{(y-1)}{6} \right) \\
&= \frac{y+1}{3}.
\end{aligned}
$$

$$
\begin{aligned}
E(y \mid x) &= \sum_{y=x+1}^{N} \frac{2y(y-x)}{(n-x)(n-x+1)} \\
&= \frac{2}{(n-x)(n-x+1)} \sum_{y=x+1}^{N} y^2 - yx \\
&= \frac{2}{(n-x)(n-x+1)} \left(\frac{n^3-x^3}{3} + \frac{n^2-x^2}{2} + \frac{n-x}{6} - \frac{nx(n+1)}{2} + \frac{x^2(x+1)}{2} \right) \\
&= \frac{2N+1+x}{3}.
\end{aligned}
$$

Ejemplo 3.3.2 (función masa de probabilidad condicional y esperanza condicional discreta)
Sea

$$f_{X,Y}(x,y) = e^{-y}\left(1-e^{-x}\right) I_{(0,y)}(x) I_{(0,\infty)}(y) + e^{-x}\left(1-e^{-y}\right) I_{(0,x)}(y) I_{(0,\infty)}(x)$$

a) Demostrar que $f_{X,Y}(x,y)$ es $fda\ conjunta$.
b)Calcular $f(y \mid x)$ y $E(y \mid x)$ para $x > 0$.

a)Una $fda\ conjunta$ cumple con $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dxdy = 1$, para este caso

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dxdy = \int_0^{\infty}\int_0^{y} e^{-y}\left(1-e^{-x}\right) dxdy + \int_0^{\infty}\int_0^{x} e^{-x}\left(1-e^{-y}\right) dydx$$

$$= \int_0^{\infty}\int_0^{y} e^{-y} dxdy - \int_0^{\infty}\int_0^{y} e^{-y}e^{-x} dxdy + \int_0^{\infty}\int_0^{x} e^{-x} dydx - \int_0^{\infty}\int_0^{x} e^{-x}e^{-y} dydx$$

$$= \left[-ye^{-y} - e^{-y} - \frac{(e^{-y}-1)}{2}\right]\Bigg|_{y=0}^{y=\infty} + \left[-xe^{-x} - e^{-x} - \frac{(e^{-x}-1)}{2}\right]\Bigg|_{x=0}^{x=\infty}$$

$$= \frac{1}{2} + \frac{1}{2} = 1.$$

b)Por propiedades de la $función\, indicadora$ (1.1.5)

$$f_{X,Y}(x,y) = \begin{cases} e^{-y}(1-e^{-x}) & 0 < x < y < \infty \\ e^{-x}(1-e^{-y}) & 0 < y < x < \infty \\ 0 & c.o.c \end{cases}$$

Se calcula $f_X(x)$ para el caso $0 < x < y < \infty$

$$\begin{aligned} f_X(x) &= \int_x^{\infty} e^{-y}\left(1-e^{-x}\right) dy \\ &= e^{-x}\left(1-e^{-x}\right) \end{aligned}$$

Y para $0 < y < x < \infty$

$$\begin{aligned} f_X(x) &= \int_0^{x} e^{-x}\left(1-e^{-y}\right) dy \\ &= e^{-x}\left(e^{-x}+x-1\right) \end{aligned}$$

Se procede a calcular $f(y \mid x)$ y $E(y \mid x)$

$$f(y \mid x) = \begin{cases} \frac{e^{-y}\left(1-e^{-x}\right)}{e^{-x}(1-e^{-x})} = \frac{e^{-y}}{e^{-x}} & 0 < x < y < \infty \\ \frac{e^{-x}\left(1-e^{-y}\right)}{e^{-x}(e^{-x}+x-1)} = \frac{1-e^{-y}}{e^{-x}+x-1} & 0 < y < x < \infty \end{cases}$$

Al aplicar la definición de esperanza condicional (2.3.2) se debe obtener

$$\begin{aligned}\int_x^\infty y\frac{e^{-y}}{e^{-x}}dy &= -e^x\left(ye^{-y}+e^{-y}\big|_x^\infty\right)\\ &= x+1\end{aligned}$$

$$\begin{aligned}\int_0^x y\frac{1-e^{-y}}{e^{-x}+x-1}dy &= \int_0^x \frac{y}{e^{-x}+x-1}-\frac{ye^{-y}}{e^{-x}+x-1}dy\\ &= \frac{x^2+2xe^{-x}+2e^{-x}-2}{2\left(e^{-x}+x-1\right)}\end{aligned}$$

$$E\left(y\mid x\right)=\begin{cases}x+1 & 0<x<y<\infty\\ \frac{x^2+2xe^{-x}+2e^{-x}-2}{2(e^{-x}+x-1)} & 0<y<x<\infty\end{cases}.$$

3.4. Transformaciones multidimensionales

Ejemplo 3.4.1 (Transformación discreta) Sean X_1 y X_2 dos variables aleatorias independientes con distribución Binomial con parámetros (n_1,p) y (n_2,p) respectivamente. Obtener la fmp de la transformación $U=X_1+X_2$.

Las fmp marginales de X_1 y X_2 son:

$$\begin{aligned}f_{X_1}(x_1) &= \binom{n_1}{x_1}p^{x_1}(1-p)^{n_1-x_1}\quad x_1=0,\,1,\,2,...n;\;\; 0\le p\le 1\\ f_{X_2}(x_2) &= \binom{n_2}{x_2}p^{x_2}(1-p)^{n_2-x_2}\quad x_2=0,\,1,\,2,...n;\;\; 0\le p\le 1\end{aligned}$$

Como son variables aleatorias independientes, su $fmp conjunta$ de (X_1,X_2) está dada por la multiplicación de las fmp respectivas de X_1 y X_2. Por lo tanto:

$$f_{X_1,X_2}(x_1,x_2)=\binom{n_1}{x_1}\binom{n_2}{x_2}p^{x_1+x_2}(1-p)^{n_1+n_2-x_1-x_2}\quad x_1,x_2=0,\,1,\,2,\ldots n;\;\; 0\le p\le 1$$

El conjunto $\mathscr{A}$ es $\{(x_1,x_2):x_1=0,\,1,\,2,...n\;y\;x_2=0,\,1,\,2,...n\}$. Se necesita la $fmp conjunta$ de $U=X_1+X_2$, se crea una variable auxiliar $V=X_2$.
Por lo tanto se tiene:

$$\Phi_1(x_1,x_2)=x_1+x_2\;\;y\;\;\Phi_2(x_1,x_2)=x_2.$$

La variable $v=x_2$ otorga a v valores no negativos, y para un valor de v fijo
$u=x_1+x_2=x_1+v$, u debe de ser un entero mayor o igual a v. Entonces el conjunto de posibles valores de (u,v) es el conjunto

$\mathscr{B} = \{(u,v):\ v = 0,\, 1,\, 2, \ldots n \ y \ u = v,\, v+1,\, v+2, \ldots\}$
Los únicos valores que cumplen con el sistema

$$\begin{aligned} x_1 + x_2 &= u \\ x_2 &= v \end{aligned}$$

Son:

$$\begin{aligned} x_1 &= u - v \\ x_2 &= v \end{aligned}$$

Por lo tanto la $fmp conjunta$ de (U,V) es:

$$\begin{aligned} f_{(U,V)}(u,v) &= f_{X_1,X_2}(u-v,v) \\ &= \binom{n_1}{u-v}\binom{n_2}{v} p^u (1-p)^{n_1+n_2-u} \qquad \begin{matrix} v = 0,\, 1,\, 2,\ n \\ u = v,\, v+1,\, v+2, \ldots \end{matrix} \end{aligned}$$

Ya obtenida la $fmp conjunta$ ahora es necesario obtener $fmp marginal$ de U. Para algún entero no negativo u, $f_{(U,V)}(u,v) > 0$ soló para $v = 0,\, 1,\, \ldots, u$. De está forma

$$\begin{aligned} f_U(u) &= \sum_{v=0}^{u} \binom{n_1}{u-v}\binom{n_2}{v} p^u (1-p)^{n_1+n_2-u} \\ &= p^u (1-p)^{n_1+n_2-u} \sum_{v=0}^{u} \binom{n_1}{u-v}\binom{n_2}{v} \end{aligned}$$

$$\left(como \sum_{w=0}^{k} \binom{n_1}{w}\binom{n_2}{k-w} = \binom{n_1+n_2}{k} \right)$$

$$= p^u (1-p)^{n_1+n_2-u} \binom{n_1+n_2}{u}$$

Por lo tanto

$$f_U(u) = p^u (1-p)^{n_1+n_2-u} \binom{n_1+n_2}{u}$$

Con lo que la transformación $U = X_1 + X_2$ se distribuye como una variable aleatoria Binomial con parámetros $(n_1 + n_2, p)$.

Ejemplo 3.4.2 (Proceso Poisson)
Al partir del supuesto de que ciertos eventos son modelados bajo un proceso Poisson con su respectivo parámetro de ocurrencia en el tiempo λ. La variable aleatoria Y representará el tiempo que tardan en ocurrir los eventos, por lo tanto $Y \sim Poisson(\lambda)$. Sea T que representa la longitud de tiempo hasta que hay n llegadas. Dado que el tiempo es continuo, se tiene que T es una variable aleatoria continua, el interés central es el de encontrar la fdp de T.

Lo que T representa, es que la probabilidad de observar la n-esima llegada después del tiempo t será la misma que la probabilidad de que se observen menos de n llegadas desde el tiempo cero hasta el tiempo t. La función de distribución acumulativa de T está dada por:

$$\begin{aligned} F(t) &= P(T \leq t) \\ &= 1 - P(T > t) \\ &= 1 - P(Y < n) \\ &= 1 - P(Y \leq n-1) \end{aligned}$$

Y tiene distribución Poisson con parámetro λ y tendrá parámetro λt en el intervalo de tiempo $(0,t)$. Por lo tanto

$$\begin{aligned} F(t) &= 1 - P(Y \leq n-1) \\ &= 1 - \sum_{y=0}^{n-1} \frac{(\lambda t)^y}{y!} e^{-\lambda t} \\ &= 1 - e^{-\lambda t} \sum_{y=0}^{n-1} \frac{(\lambda t)^y}{y!}. \end{aligned}$$

Por el teorema fundamental del cálculo se sabe que $f(t) = F'(t)$, al usar la regla de la cadena, la fdp vendrá dada por:

$$\begin{aligned} f(t) &= \lambda e^{-\lambda t} \sum_{y=0}^{n-1} \frac{(\lambda t)^y}{y!} - e^{-\lambda t} \lambda \sum_{y=0}^{n-1} \frac{y \cdot (\lambda t)^{y-1}}{y!} \\ &= \lambda e^{-\lambda t} \left[\sum_{y=0}^{n-1} \frac{(\lambda t)^y}{y!} - \sum_{y=0}^{n-1} \frac{y \cdot (\lambda t)^{y-1}}{y!} \right] \\ &= \lambda e^{-\lambda t} \left[\sum_{y=0}^{n-1} \frac{(\lambda t)^y}{y!} - \sum_{y=1}^{n-1} \frac{(\lambda t)^{y-1}}{(y-1)!} \right] \\ &= \lambda e^{-\lambda t} \left[\sum_{y=0}^{n-1} \frac{(\lambda t)^y}{y!} - \sum_{y=0}^{n-2} \frac{(\lambda t)^{y-1}}{(y-1)!} \right] \end{aligned}$$

$(al\ hacer\ x = y-1)$

$$\begin{aligned} &= \lambda e^{-\lambda t} \left[\sum_{y=0}^{n-1} \frac{(\lambda t)^y}{y!} - \sum_{x=0}^{n-2} \frac{(\lambda t)^x}{x!} \right] \\ &= \lambda e^{-\lambda t} \left[\frac{(\lambda t)^{n-1}}{(n-1)!} + \sum_{y=0}^{n-2} \frac{(\lambda t)^y}{y!} - \sum_{x=0}^{n-2} \frac{(\lambda t)^x}{x!} \right] \\ &= \lambda e^{-\lambda t} \left[\frac{(\lambda t)^{n-1}}{(n-1)!} \right] \\ &= \frac{t^{n-1} \lambda^n e^{-\lambda t}}{(n-1)!}. \end{aligned}$$

Como $\Gamma(n) = (n-1)!$ y como esto es valido ya que n es entero

$$f(t) = \frac{t^{n-1}\lambda^n e^{-\lambda t}}{\Gamma(n)}.$$

Si se supone que $t = x$, $n = \alpha$ y $\beta = \frac{1}{\lambda}$ entonces se puede reescribir $f(t)$ de la siguiente manera

$$f(x) = \frac{x^{\alpha-1}e^{-\frac{x}{\beta}}}{\Gamma(\alpha)\beta^{\alpha}}$$

como $t \geq 0 \Rightarrow x \geq 0$, $n > 0 \Rightarrow \alpha > 0$ y por último $\lambda > 0 \Rightarrow \beta > 0$.

Finalmente se observa que $f(t)$ es un caso especial de la distribución Gamma cuando sus parámetros son $n = \alpha$ y $\beta = \frac{1}{\lambda}$.

Si $Y \sim Poisson(\lambda)$, el tiempo hasta que se tengan n llegadas y que se representa por la variable aleatoria T será descrito por una distribución Gamma, esto es

$$T \sim Gamma(n, \frac{1}{\lambda}).$$

Ejemplo 3.4.3 (Convolucion) Calcular la distribución de la la suma de dos variables aleatorias X y Y independientes cada una con distribución $Cauchy(0,1)$.

La fdp de una variable aleatoria $Cauchy(\theta,\sigma)$ es de la forma

$$f_X(x,\theta,\sigma) = \frac{1}{\pi\sigma}\frac{1}{1+\left(\frac{x-\theta}{\sigma}\right)^2}, \ -\infty < x < \infty, \ -\infty < \theta < \infty, \ \sigma > 0.$$

Entonces

$$f_X(x,0,1) = \frac{1}{\pi}\frac{1}{1+x^2}, \ \ f_Y(y,0,1) = \frac{1}{\pi}\frac{1}{1+y^2}, \ -\infty < x < \infty, \ -\infty < y < \infty$$

Al definir la nueva variable aleatoria $Z = X + Y$ y por el teorema de convoluciones (2.4.3) y con $W = X$ se tiene

$$\begin{aligned} f_Z(z) &= \int_{-\infty}^{\infty} f_X(w) f_Y(z-w) dw \\ &= \int_{-\infty}^{\infty} \frac{1}{\pi}\frac{1}{1+w^2}\frac{1}{\pi}\frac{1}{1+(z-w)^2} dw \\ &= \frac{1}{\pi^2}\int_{-\infty}^{\infty} \frac{1}{1+w^2}\frac{1}{1+(z-w)^2} dw \end{aligned}$$

Se procede a resolver la integral por el método de fracciones parciales[4] con lo que

$$\frac{1}{1+(z-w)^2}\frac{1}{1+w^2} = \frac{A(z-w)+B}{1+(z-w)^2} + \frac{Cw+D}{1+w^2} \Rightarrow$$

[4]George B. Thomas, Jr. (2006). Cálculo. Una variable. Undécima edición. México. Pearson. (p.570)

$$1=(A(z-w)+B)\left(1+w^2\right)+(Cw+D)\left(1+(z-w)^2\right),$$

Se buscan los valores de $A, B, C, y D$ que dependen de z y no de w y se obtiene el sistema de ecuaciones

$$\begin{aligned}
Az+B+D+Dz^2 &= 1\\
C+Cz^2-A-D2z &= 0\\
Az+B-C2z+D &= 0\\
-A+C &= 0.
\end{aligned}$$

Al resolver el sistema se obtiene

$$A=C=\frac{2}{z(z^2+4)}\ y\,D=B=\frac{z}{z(z^2+4)},$$

Al continuar con la integral

$$\begin{aligned}
f_Z(z) &= \frac{1}{\pi^2}\int_{-\infty}^{\infty}\left[\frac{1}{1+w^2}\frac{1}{1+(z-w)^2}\right]dw\\
&= \frac{1}{\pi^2}\int_{-\infty}^{\infty}\left[\frac{A(z-w)+B}{1+(z-w)^2}+\frac{Cw+D}{1+w^2}\right]dw\\
&= \frac{1}{\pi^2}\int_{-\infty}^{\infty}\left[\frac{\frac{2}{z(z^2+4)}(z-w)}{1+(z-w)^2}+\frac{\frac{z}{z(z^2+4)}}{1+(z-w)^2}+\frac{\frac{2}{z(z^2+4)}w}{1+w^2}+\frac{\frac{z}{z(z^2+4)}}{1+w^2}\right]dw
\end{aligned}$$

Ahora se agrupan estas integrales

$$\begin{aligned}
f_Z(z)=&\left\{\frac{1}{\pi^2(z^2+4)}\int_{-\infty}^{\infty}\left[\frac{2}{1+(z-w)^2}+\frac{1}{1+(z-w)^2}+\frac{1}{1+w^2}\right]dw\right\}+\\
&+\left\{\frac{1}{\pi^2 z(z^2+4)}\int_{-\infty}^{\infty}\left[\frac{2w}{1+w^2}-\frac{2w}{1+(z-w)^2}\right]dw\right\}
\end{aligned}$$

Para la primer integral se tiene

$$\begin{aligned}
&\frac{1}{\pi^2(z^2+4)}\int_{-\infty}^{\infty}\left[\frac{2}{1+(z-w)^2}+\frac{1}{1+(z-w)^2}+\frac{1}{1+w^2}\right]dw=\\
&\frac{-3}{\pi^2(z^2+4)}\left[arctan(z-w)\right]_{-\infty}^{\infty}+\frac{1}{\pi^2(z^2+4)}\left[arctan(w)\right]_{-\infty}^{\infty}=\\
&\frac{3}{\pi(z^2+4)}+\frac{1}{\pi(z^2+4)}=\\
&\frac{4}{\pi(z^2+4)}
\end{aligned}$$

Mientras que para la segunda

$$\frac{1}{\pi^2 z(z^2+4)}\int_{-\infty}^{\infty}\left[\frac{2w}{1+w^2}-\frac{2w}{1+(z-w)^2}\right]dw =$$

$$\frac{2}{\pi^2 z(z^2+4)}\int_{-\infty}^{\infty}\left[\frac{w}{1+w^2}-\frac{w}{1+(z-w)^2}\right]dw =$$

$$\frac{2}{\pi^2 z(z^2+4)}\int_{-\infty}^{\infty}\left[\frac{w}{1+w^2}-\frac{(w-z)}{1+(w-z)^2}\right]dw - \frac{2z}{\pi^2 z(z^2+4)}\int_{-\infty}^{\infty}\frac{1}{1+(w-z)^2}dw =$$

$$\frac{2}{\pi^2 z(z^2+4)}\left(\left[\frac{1}{2}log(1+w^2)\right]_{-\infty}^{\infty}-\left[\frac{1}{2}log(1+(w-z)^2)\right]_{-\infty}^{\infty}\right)-\frac{2z}{\pi^2 z(z^2+4)}\left[arctan(w-z)\right]_{-\infty}^{\infty}$$

$$= \frac{-2}{\pi(z^2+4)}.$$

Es importante notar el hecho de que la integral

$$\int_{-\infty}^{\infty}\left[\frac{w}{1+w^2}-\frac{(w-z)}{1+(w-z)^2}\right]dw = \left[\frac{1}{2}log(1+w^2)\right]_{-\infty}^{\infty}-\left[\frac{1}{2}log(1+(w-z)^2)\right]_{-\infty}^{\infty},$$

es finita e igual a cero.

Al sumar ambas integrales se obtiene

$$\begin{aligned} f_Z(z) &= \frac{4}{\pi(z^2+4)}-\frac{2}{\pi(z^2+4)} \\ &= \frac{2}{\pi(z^2+4)} \\ &= \frac{1}{2\pi}\frac{4}{(z^2+4)} \\ &= \frac{1}{2\pi}\frac{1}{1+\left(\frac{z}{2}\right)^2}. \end{aligned}$$

Por lo tanto la fdp de la variable aleatoria Z es

$$f_Z(z) = \frac{2}{\pi(z^2+4)}$$

Con lo que Z tiene una distribución $Cauchy(0,2)$.

3.5. Estadísticos de orden

Ejemplo 3.5.1 (Fdp estadísticos de orden) Sean $X_1,\ X_2,\ X_3$ tres variables aleatorias con misma distribución $Weibull(2,\beta)$. Encontrar la probabilidad de que la distancia entre cada par de ellas sea mayor a 1.
Por lo tanto se necesita
$P\left[|X_2 - X_1| > 1, |X_3 - X_1| > 1,\ |X_3 - X_2| > 1\right]$

Sean $X_{(1)},\ X_{(2)},\ X_{(3)}$ los estadísticos de orden de $X_1,\ X_2,\ X_3$ respectivamente, por (2.5.1) se tiene que la $fdp\ conjunta$ de $X_{(1)},\ X_{(2)},\ X_{(3)}$ es:

$$f_{X_{(1)},\ X_{(2)},\ X_{(3)}}(x_1,\ x_2,\ x_3) = \begin{cases} 3!\frac{2^3}{\beta^3}x_1e^{-x_1^2/\beta}x_2e^{-x_2^2/\beta}x_3e^{-x_3^2/\beta} & 0 < x_1 < x_2 < x_3 \\ 0 & c.\ o.\ c. \end{cases}$$

Con la $fdp\ conjunta$ de $X_{(1)},\ X_{(2)},\ X_{(3)}$ se puede obtener la probabilidad deseada de la siguiente manera:

$$\begin{aligned} P\left[|X_2 - X_1| > 1, |X_3 - X_1| > 1,\ |X_3 - X_2| > 1\right] &= \\ P\left[X_{(2)} > X_{(1)} + 1, X_{(3)} > X_{(2)} + 1\right] &= \\ \int_0^\infty \int_{x_1+1}^\infty \int_{x_2+1}^\infty f_{X_{(1)},\ X_{(2)},\ X_{(3)}}(x_1,\ x_2,\ x_3)\,dx_3dx_2dx_1 &= \\ \int_0^\infty \int_{x_1+1}^\infty \int_{x_2+1}^\infty 3!\frac{2^3}{\beta^3}x_1e^{-x_1^2/\beta}x_2e^{-x_2^2/\beta}x_3e^{-x_3^2/\beta}dx_3dx_2dx_1 &= \\ \int_0^\infty \int_{x_1+1}^\infty \left[-\frac{24}{\beta^2}x_1e^{-x_1^2/\beta}x_2e^{-x_2^2/\beta}\int_{x_2+1}^\infty -\frac{2}{\beta}x_3e^{-x_3^2/\beta}dx_3\right]dx_2dx_1 &= \\ \int_0^\infty \int_{x_1+1}^\infty \frac{24}{\beta^2}x_1x_2e^{-\left(\frac{x_1^2}{\beta}+\frac{x_2^2}{\beta}+\frac{(x_2+1)^2}{\beta}\right)}dx_2dx_1 &= \\ \int_0^\infty \left[\frac{6\cdot 2^{\frac{3}{2}}}{\beta^{\frac{3}{2}}}x_1e^{\frac{-1}{\beta}\left(x_1^2+\frac{1}{2}\right)}\int_{x_1+1}^\infty e^{-\left(\frac{\sqrt{2}}{\sqrt{\beta}}\left(x_2+\frac{1}{2}\right)\right)^2}\frac{\sqrt{2}}{\sqrt{\beta}}x_2dx_2\right]dx_1 & \end{aligned}$$

$\left(\textit{Al hacer el cambio de variable } u = x_2 + \frac{1}{2} \Rightarrow du = dx_2 \textit{ y } x_2 = u - \frac{1}{2}\right)$

$$= \int_0^\infty \left[\frac{6\cdot 2^{\frac{3}{2}}}{\beta^{\frac{3}{2}}}x_1e^{\frac{-1}{\beta}\left(x_1^2+\frac{1}{2}\right)}\int_{x_1+\frac{3}{2}}^\infty e^{-\left(\frac{\sqrt{2}}{\sqrt{\beta}}u\right)^2}\frac{\sqrt{2}}{\sqrt{\beta}}\left(u-\frac{1}{2}\right)du\right]dx_1$$

La integral $\int_{x_1+\frac{3}{2}}^\infty e^{-\left(\frac{\sqrt{2}}{\sqrt{\beta}}u\right)^2}\frac{\sqrt{2}}{\sqrt{\beta}}\left(-\frac{1}{2}\right)du = -\frac{1}{2}\int_{x_1+\frac{3}{2}}^\infty e^{-\left(\frac{\sqrt{2}}{\sqrt{\beta}}u\right)^2}\frac{\sqrt{2}}{\sqrt{\beta}}du$ se resuelve como la integral (1.1.4), es necesario tener cuidado de los límites de integración y del saber que $A \neq 1$, después de
operaciones algebraicas se obtiene

$$\int_0^{\infty}\left[\frac{6\cdot 2^{\frac{3}{2}}}{\beta^{\frac{3}{2}}}x_1e^{\frac{-1}{\beta}\left(x_1^2+\frac{1}{2}\right)}\int_{x_1+\frac{3}{2}}^{\infty}e^{-\left(\frac{\sqrt{2}}{\sqrt{\beta}}u\right)^2}\frac{\sqrt{2}}{\sqrt{\beta}}\left(u-\frac{1}{2}\right)du\right]dx_1=$$

$$\int_0^{\infty}\left[\frac{6\cdot 2^{\frac{3}{2}}}{\beta^{\frac{3}{2}}}x_1e^{\frac{-1}{\beta}\left(x_1^2+\frac{1}{2}\right)}\left(\frac{\sqrt{\beta}}{2^{\frac{3}{2}}}e^{\frac{-2}{\beta}\left(x_1+\frac{3}{2}\right)^2}-\frac{\sqrt{\pi}}{4}e^{\frac{-2}{\beta}\left(x_1+\frac{3}{2}\right)^2}\right)\right]dx_1=$$

$$\int_0^{\infty}\left[\frac{6\cdot 2^{\frac{3}{2}}}{\beta^{\frac{3}{2}}}x_1e^{\frac{-1}{\beta}\left(x_1^2+\frac{1}{2}\right)}e^{\frac{-2}{\beta}\left(x_1+\frac{3}{2}\right)^2}\left(\frac{\sqrt{\beta}}{2^{\frac{3}{2}}}-\frac{\sqrt{\pi}}{4}\right)\right]dx_1=$$

$$\frac{6\cdot 2^{\frac{3}{2}}}{\beta^{\frac{3}{2}}}\left(\frac{\sqrt{\beta}}{2^{\frac{3}{2}}}-\frac{\sqrt{\pi}}{4}\right)e^{\frac{-2}{\beta}}\int_0^{\infty}e^{\frac{-3}{\beta}(x_1+1)^2}x_1dx_1=$$

$(al\ \ hacer\ \ el\ \ cambio\ \ de\ \ variable\ \ u=x_1+1\Rightarrow du=dx_1\ \ y\ \ x_1=u-1)$

$$\frac{6\cdot 2^{\frac{3}{2}}}{\beta^{\frac{3}{2}}}\left(\frac{\sqrt{\beta}}{2^{\frac{3}{2}}}-\frac{\sqrt{\pi}}{4}\right)e^{\frac{-2}{\beta}}\int_1^{\infty}e^{\frac{-3}{\beta}u^2}(u-1)\,du$$

La integral $\int_1^{\infty}-e^{\frac{-3}{\beta}u^2}du=-\frac{\sqrt{\beta}}{\sqrt{3}}\int_1^{\infty}e^{-\left(\frac{\sqrt{3}}{\sqrt{\beta}}u\right)^2}\frac{\sqrt{3}}{\sqrt{\beta}}du$, se transforma a

$$\frac{6\cdot 2^{\frac{3}{2}}}{\beta^{\frac{3}{2}}}\left(\frac{\sqrt{\beta}}{2^{\frac{3}{2}}}-\frac{\sqrt{\pi}}{4}\right)e^{\frac{-2}{\beta}}\int_1^{\infty}e^{\frac{-3}{\beta}u^2}(u-1)\,du=$$

$$\frac{6\cdot 2^{\frac{3}{2}}}{\beta^{\frac{3}{2}}}\left(\frac{\sqrt{\beta}}{2^{\frac{3}{2}}}-\frac{\sqrt{\pi}}{4}\right)e^{\frac{-2}{\beta}}\left(e^{\frac{-3}{\beta}}\left(\frac{\beta}{6}-\frac{\sqrt{\pi\beta}}{2\sqrt{3}}\right)\right)=$$

$$e^{\frac{-5}{\beta}}\left(\frac{\beta\sqrt{6}-\sqrt{18\pi\beta}+3\pi+\sqrt{3\pi\beta}}{\beta\sqrt{6}}\right)$$

Con lo que

$$P\left[X_{(2)}>X_{(1)}+1,X_{(3)}>X_{(2)}+1\right]=e^{\frac{-5}{\beta}}\left(\frac{\beta\sqrt{6}-\sqrt{18\pi\beta}+3\pi+\sqrt{3\pi\beta}}{\beta\sqrt{6}}\right)\ \ \beta>0.$$

Ejemplo 3.5.2 (Transformación estadísticos de orden) Sean X_1 y X_2 dos variables aleatorias independientes con distribución exponencial de parámetro λ. Demostrar que las variables aleatorias $U=mín(X_1,X_2)$ y $V=máx(X_1,X_2)-mín(X_1,X_2)$ son independientes.

Sean $X_{(1)},X_{(2)}$ los estadísticos de orden de las variables aleatorias X_1,X_2, por (2.5.2) la $fdp\ conjunta$ de $X_{(1)},X_{(2)}$ es de la forma

$$f_{X_{(1)},X_{(2)}}(x_1,x_2)=2\frac{e^{\frac{-x_1}{\lambda}}}{\lambda}\frac{e^{\frac{-x_2}{\lambda}}}{\lambda}\quad 0<x_1<x_2<\infty$$

Sea $u=x_{(1)}$ y $v=x_{(2)}-x_{(1)} \Rightarrow x_{(1)}=u$ y $x_{(2)}=v+u$, como $u,v>0 \Rightarrow 0<u<u+v<\infty$.

$$J_{(u,v)}=\begin{vmatrix} 1 & 1 \\ 0 & 1 \end{vmatrix}=|1|=1$$

Por lo tanto la $fdp\ conjunta$ de U y V es

$$f_{U,\,V}(u,\ v)=2\frac{e^{\frac{-u}{\lambda}}}{\lambda}\frac{e^{\frac{-(v+u)}{\lambda}}}{\lambda}\quad 0<u<u+v<\infty$$

$$=\frac{2e^{\frac{-2u}{\lambda}}}{\lambda}\frac{e^{\frac{-v}{\lambda}}}{\lambda}\quad \begin{cases}0<v<\infty\\ 0<u<\infty\end{cases}$$

Se observa que U se distribuye como una variable aleatoria exponencial con parámetro $\frac{2}{\lambda}$ y V como una variable aleatoria exponencial con parámetro $\frac{1}{\lambda}$. Por lo tanto U y V son variables aleatorias independientes.

Ejemplo 3.5.3 (Fdp marginales estadísticos de orden) Sean $X_1,\ X_2,\ldots,\ X_n$ variables aleatorias independientes con distribución $Uniforme(0,\ 1)$. Encontrar

a) La fdp de la transformación $U=X_{(n)}-X_{(1)}$.
b) La fdp de la transformación $V=\frac{X_{(n)}+X_{(1)}}{2}$.
a) La $fdp\ conjunta$ de $(U,\ V)$.

Sean $X_{(1)},\ X_{(2)},\ldots,\ X_{(n)}$ los estadísticos de orden de $X_1,\ X_2,\ldots,\ X_3$ respectivamente, entonces la $fdp\ conjunta$ de $X_{(1)},\ X_{(n)}$ es:

$$f_{X_{(1)},\,X_{(n)}}(x_1,\ x_n)=\begin{cases}n(n-1)\left[x_n-x_1\right]^{n-2} & 0<x_1<x_n<1\\ 0 & c.\ o.\ c.\end{cases}$$

a) Sea $U=X_{(n)}-X_{(1)}$ y $W=X_{(1)} \Rightarrow X_{(1)}=W$ y $X_{(n)}=U+W$

$$J_{\left(X_{(1)},\,X_{(n)}\right)}=\begin{vmatrix} 1 & 1 \\ 0 & 1 \end{vmatrix}=1$$

Como $0<x_1<x_n<1 \Rightarrow 0<w<u+w<1$, la $fdp\ conjunta$ de (U,W) es de la forma

$$f_{U,\,W}(u,\ w)=n(n-1)\,u^{n-2}\qquad 0<w<u+w<1$$

Se obtiene la fdp marginal de U, como $0<w$ y $u+w<1 \Rightarrow 0<w<1-u$ se tiene

$$\begin{aligned} f_U(u) &= \int_0^{1-u} n(n-1)u^{n-2}dw \\ &= (1-u)n(n-1)u^{n-2} \;\; 0<u<1 \end{aligned}$$

b) Sea $V = \frac{X_{(n)}+X_{(1)}}{2}$ y $W = X_{(1)} \Rightarrow X_{(1)} = W$ y $X_{(n)} = 2V - W$

$$J_{\left(X_{(1)},\, X_{(n)}\right)} = \begin{vmatrix} 2 & 0 \\ -1 & 1 \end{vmatrix} = 2$$

Al tener la condición $0 < x_1 < x_n < 1 \Rightarrow 0 < w < 2v - w < 1$, por lo tanto la $fdp\ conjunta$ de (V, W) es

$$f_{V,\,W}(v,\ w) = 2n(n-1)(2v-2w)^{n-2} \qquad 0 < w < 2v-w < 1$$

Se tiene la región de integración delimitada por las condiciones $0 < w$, $w < 2v - w$ y $2v - w < 1$, al dejar libre a la variable aleatoria V entre 0 y 1 se tiene que

$$\begin{array}{lll} si & 0 < v < \frac{1}{2} & 0 < w < v \\ si & \frac{1}{2} \le v < 1 & 2v-1 < w < v \end{array}$$ (figura 3.5.1), por lo tanto la fdp de V es

$$\begin{aligned} f_V(v) &= \int_0^{v} 2^{n-1} n(n-1)(v-w)\,dw \\ &= n2^{n-1}v^{n-1} \qquad 0 < v < \frac{1}{2} \end{aligned}$$

$$\begin{aligned} f_V(v) &= \int_{2v-1}^{v} 2^{n-1} n(n-1)(v-w)\,dw \\ &= n2^{n-1}(1-v)^{n-1} \qquad \frac{1}{2} < v < 1. \end{aligned}$$

c) Sea $U = X_{(n)} - X_{(1)}$ y $V = \frac{X_{(n)}+X_{(1)}}{2} \Rightarrow X_{(n)} = V + \frac{U}{2}$ y $X_{(1)} = V - \frac{U}{2}$

$$J_{\left(X_{(1)},\, X_{(n)}\right)} = \begin{vmatrix} 1 & 1 \\ \frac{1}{2} & -\frac{1}{2} \end{vmatrix} = 1$$

Con lo que la $fdp\ conjunta$ de $(U,\ V)$ es

$$\begin{aligned} f_{U,\,V}(u,\ v) &= n(n-1)\left(v + \frac{u}{2} - v + \frac{u}{2}\right)^{n-2} \\ &= n(n-1)u^{n-2} \;\; 0 < v - \frac{u}{2} < v + \frac{u}{2} < 1. \end{aligned}$$

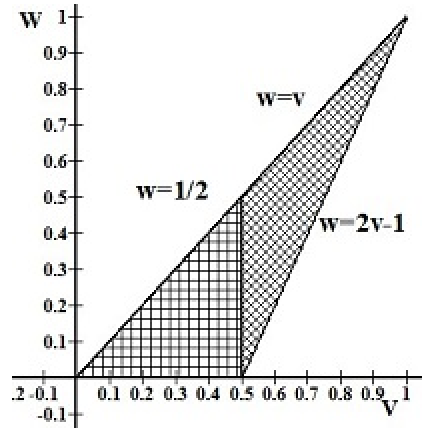

FIGURA 3.5.1. Región de integración ejemplo 3.5.3.

3.6. Máxima verosimilitud.

Ejemplo 3.6.1 (Máxima verosimilitud de una muestra aleatoria discreta) Sean $X_1, X_2, \ldots, X_n$ variables aleatorias que provienen de una muestra aleatoria $Geometrica(p)$. Determinar el estimador máximo verosímil de p.
La fmp de una variable aleatoria $Geometrica(p)$ es:

$$f_X(x;p) = (1-p)^{x-1}\, p, \;\; x = 1,2,...;\; 0 \leq p \leq 1.$$

De (2.6.1) la función de verosimilitud es

$$\begin{aligned} L(p,\underline{X}) =& \prod_{i=1}^{n} f_X(x_i,p) \\ =& \prod_{i=1}^{n} (1-p)^{x-1}\, p \\ =& \; p^n (1-p)^{-n} (1-p)^{\sum_{i=1}^{n} x_i} \end{aligned}$$

Ahora se obtiene la log-verosimilitud

$$logL(p,\underline{X}) = \ell(p,\underline{X}) = nlog(p) + (log(1-p))\left(\sum_{i=1}^{n} x_i - n\right)$$

Al derivar e igualar a cero, se obtiene

$$\frac{\delta \ell(p,\underline{X})}{\delta p} = \frac{n}{p} - \frac{\sum_{i=1}^{n} x_i - n}{1-p} = 0 \implies$$

$$\hat{p} = \frac{n}{\sum_{i=1}^{n} x_i}$$

Es momento de verificar sí es máximo

$$\frac{\delta^2 \ell(p,\underline{X})}{\delta^2 p} = -\frac{n}{p^2} - \frac{\sum\limits_{i=1}^{n} x_i - n}{(1-p)^2} \,|_{\hat{p}} = -\frac{n}{\frac{n}{\sum\limits_{i=1}^{n} x_i}^2} - \frac{\sum\limits_{i=1}^{n} x_i - n}{\left(1 - \frac{n}{\sum\limits_{i=1}^{n} x_i}\right)^2} < 0.$$

Ya que $\sum\limits_{i=1}^{n} x_i - n \geq 0$, $\hat{p} = \frac{n}{\sum\limits_{i=1}^{n} x_i}$ es el estimador máximo verosímil de p.

Ejemplo 3.6.2 (Máxima verosimilitud de una muestra aleatoria continua) Sean $X_1, X_2, \ldots, X_n$ variables aleatorias que provienen de una muestra aleatoria $Normal(\mu, \sigma^2)$. Encontrar el estimador máximo verosímil para μ y σ^2.

La fdp de cada variable aleatoria X_i se distribuye $Normal(\mu, \sigma^2)$ y está dada por

$$f(x;\mu,\sigma^2) = \frac{1}{\sqrt{2\pi}\sigma} e^{\frac{-(x-\mu)^2}{2\sigma^2}}, \;\; -\infty < x < \infty, \; -\infty < \mu < \infty, \; \sigma > 0$$

La verosimilitud viene dada por

$$\begin{aligned} L(\mu,\sigma^2,\underline{X}) &= \prod_{i=1}^{n} f_X(x_i,\mu,\sigma^2) \\ &= \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}\sigma} exp\left\{\frac{-(x_i-\mu)^2}{2\sigma^2}\right\} \\ &= \left(\frac{1}{\sqrt{2\pi}\sigma}\right)^n exp\left\{\frac{-\sum\limits_{i=1}^{n}(x_i-\mu)^2}{2\sigma^2}\right\}. \end{aligned}$$

La log-verosimilitud es

$$logL(\mu,\sigma^2,\underline{X}) = \ell(\mu,\sigma^2,\underline{X}) = nlog\left(\frac{1}{\sqrt{2\pi}\sigma}\right) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2$$

Para μ se obtiene su derivada y se iguala a cero

$$\frac{\delta\ell(\mu,\sigma^2,\underline{X})}{\delta\mu} = \frac{1}{\sigma^2}\sum_{i=1}^{n}(x_i-\mu) = 0 \Rightarrow$$

$$\frac{1}{\sigma^2}\sum_{i=1}^{n} x_i = \frac{n\mu}{\sigma^2} \Rightarrow \hat{\mu} = \frac{\sum\limits_{i=1}^{n} x_i}{n} = \bar{x}.$$

Se hace lo mismo para σ^2

$$\frac{\delta\ell(\mu,\sigma^2,\underline{X})}{\delta\sigma^2} = -\frac{n}{\sigma} + \frac{\sum\limits_{i=1}^{n}(x_i-\mu)^2}{\sigma^3} = 0 \Rightarrow$$

$$\frac{n}{\sigma}=\frac{\sum_{i=1}^{n}(\ x_i-\mu)^2}{\sigma^3}\Rightarrow n\sigma^2=\sum_{i=1}^{n}(\ x_i-\mu)^2\Rightarrow$$

$$\widehat{\sigma^2}=\frac{\sum_{i=1}^{n}(\ x_i-\mu)^2}{n}.$$

Solo falta demostrar que son máximos

$$\frac{\delta^2\ell(\mu,\sigma^2,\underline{X})}{\delta^2\mu}=\frac{-n\sigma^2}{\sigma^4}=\frac{-n}{\sigma^2}\mid_{\bar{x}}=\frac{-n}{\sigma^2}<0.$$

$$\begin{aligned}\frac{\delta^2\ell(\mu,\sigma^2,\underline{X})}{\delta^2\sigma^2}&=\frac{n}{\sigma^2}-\frac{3\sum_{i=1}^{n}(\ x_i-\mu)^2}{\sigma^4}\mid_{\widehat{\sigma^2}}\\&=\frac{n}{\frac{\sum_{i=1}^{n}(x_i-\mu)^2}{n}}-\frac{3\sum_{i=1}^{n}(\ x_i-\mu)^2}{\left(\frac{\sum_{i=1}^{n}(x_i-\mu)^2}{n}\right)^2}\\&=\frac{n^2}{\sum_{i=1}^{n}(\ x_i-\mu)^2}-\frac{3n^2\sum_{i=1}^{n}(\ x_i-\mu)^2}{\left(\sum_{i=1}^{n}(\ x_i-\mu)^2\right)^2}\\&=\frac{n^2}{\sum_{i=1}^{n}(\ x_i-\mu)^2}-\frac{3n^2}{\sum_{i=1}^{n}(\ x_i-\mu)^2}\\&=-\frac{2n^2}{\sum_{i=1}^{n}(\ x_i-\mu)^2}<0.\end{aligned}$$

Por lo tanto $\hat{\mu}=\bar{x}$ y $\widehat{\sigma^2}=\frac{\sum_{i=1}^{n}(x_i-\mu)^2}{n}$ son estimadores de máxima verosimilitud para una muestra con $fdp\ Normal(\mu,\sigma^2)$.

CAPÍTULO 4

Ejercicios aplicados

Finalmente, el último paso consiste en saber aplicar adecuadamente en casos prácticos, las herramientas proporcionadas en los capítulos anteriores. Es por ello que este último capítulo se centra en cómo se puede aplicar la teoría y técnicas de probabilidad, estadística y cálculo de una y varias variables para obtener estadísticas que son útiles para la toma de decisiones.

Es importante recordar que el objetivo al trabajar con muestras aleatorias, consiste en modelar su comportamiento y adaptarlo a una función de distribución, la cual permitirá obtener estadísticas que proporcionen la información necesaria de interés.

Sin embargo, aún después de entender las bases de teoría, resulta complicado poder aplicarla a un problema dado, por eso es de suma importancia comprender el problema que se está atacando, para así poder hacer el uso adecuado de las herramientas convenientes.

4.1. Seguros

Ejemplo 4.1.1 Función de supervivencia

Sea (x) que denota una vida a edad x, donde $x \geq 0$. La muerte de (x) puede ocurrir en cualquier edad superior que x, se puede modelar el tiempo futuro de vida de (x) como una variable aleatoria continua que será denotada por T_X.
Por lo tanto $x + T_X$ representa la edad de muerte de la variable aleatoria (x).

Sea F_X la función de distribución acumulativa de T_X tal que

$$F_X(t) = Pr[T_X \leq t].$$

Entonces $F_X(t)$ representa la probabilidad de que (x) no sobreviva más allá de la edad $x+t$, y se refiere a F_X como la distribución del tiempo de vida de (x).
Al representar a T_0 como el tiempo futuro de vida al nacimiento de un individuo x entonces se obtiene está relación

$$Pr[T_X \leq t] = Pr[T_0 \leq x+t \mid T_0 > x]. \tag{4.1.1}$$

Si se está interesado en la probabilidad de supervivencia en lugar de la de muerte,
se define una nueva función S_X como

$$S_X(t) = 1 - F_X(t) = Pr[T_X > t].$$

Por lo tanto $S_X(t)$ representa la probabilidad de que (x) sobreviva por al menos t años, y S_X será denotada como la **función de supervivencia**.

Propiedades:

- $S(0) = 1$, ya que ninguna persona comienza muerta.
- $\lim_{t\to\infty} S(t) = 0$ con lo que todas las personas mueren eventualmente.
- $S(t_1) \geq S(t_2) \Leftrightarrow t_1 \leq t_2$ por lo tanto $S(t)$ es monótona decreciente.

Se pueden establecer algunas relaciones entre estas dos funciones $S_X(t)\, y\, F_X(t)$.
Al recordar la teoría de probabilidad[1] para dos eventos A y B

$$Pr[A \mid B] = \frac{Pr[A \cap B]}{Pr[B]},$$

tal que, si se define a $[T_0 \leq x+t]$ como el evento A, y a $[T_0 > x]$ como el evento B, se puede reordenar la parte derecha de (4.1.1) para obtener

$$Pr[T_X \leq t] = \frac{Pr[x < T_0 \leq x+t]}{Pr[T_0 > x]},$$

esto es,

$$F_X(t) = \frac{F_0(x+t) - F_0(x)}{S_0(x)}$$

si se considera que $S_X(t) = 1 - F_X(t)$,

$$S_X(t) = \frac{S_0(x+t)}{S_0(x)} \tag{4.1.2}$$

Sea

$$F_0(t) = 1 - (1 - t/110)^{1/5} \quad para\, 0 \leq t \leq 110.$$

Calcular la probabilidad de
(a) la vida de un recién nacido sobrevive mas allá de 40 años,
(b) una vida de edad 25 muere antes de la edad 55, y
(c) una vida de edad 50 sobrevive más allá de la edad 75.
(d) calcular $\lim_{t\to 110} S_0(t)$.

Solución: (a) La probabilidad es

[1] Mood, A.M.; Graybill, F.A, Boes, D.C. (1974). Introduction to the theory of statistics. Nueva York: McGraw-Hill. (p.32)

$$S_0(40) = 1 - F_0(40) = (1 - 40/110)^{1/5} = .913568$$

(b)

$$F_{25}(30) = \frac{F_0(55) - F_0(25)}{1 - F_0(25)} = 0.083380$$

(c)

$$S_{50}(25) = \frac{S_0(75)}{S_0(50)} = 0.897808.$$

(d)

$$\lim_{t \to 110} S_0(t) = \lim_{t \to 110} (1 - t/110)^{1/5} =$$

$$= (1-1)^{1/5} = 0.$$

Es preciso notar que para este modelo de supervivencia 110 años es la edad máxima o límite de un individuo.

Ejemplo 4.1.2 Distribución exponencial

El tiempo de espera por la primer reclamación de choque de una mujer sigue una distribución *exponencial* con media de 2, mientras que el tiempo de espera por la primer reclamación de choque de un hombre sigue una distribución *exponencial* con media 4.

Bajo el supuesto de que las reclamaciones son independientes. ¿Cuál es la probabilidad de que la primer reclamación de una mujer se presente en el primer año y la primer reclamación de un hombre se presente en los primeros tres años?

Sea X la variable aleatoria que representa la primer reclamación por choque de una mujer y Y la variable aleatoria que representa la primer reclamación por choque de un hombre. Es necesario obtener $P(X \leq 1, Y \leq 3)$

Como X y Y son independientes por (2.2.11)

$$P(X \leq 1, Y \leq 3) = P(X \leq 1)P(Y \leq 3)$$

Dado que X tiene media 2 y Y tiene media 4

$$f_X(x) = \frac{1}{2}e^{-\frac{1}{2}x} \quad y \quad f_Y(y) = \frac{1}{4}e^{-\frac{1}{4}y}$$

$$P(X \leq 1)P(Y \leq 3) = \left(-\int_0^1 -\frac{1}{2}e^{-\frac{1}{2}x}dx\right)\left(-\int_0^3 -\frac{1}{4}e^{-\frac{1}{4}y}dy\right)$$

$$= \left(-e^{-\frac{1}{2}x}\Big|_0^1\right)\left(-e^{-\frac{1}{4}x}\Big|_0^3\right) = \left(1-e^{-\frac{1}{2}}\right)\left(1-e^{-\frac{3}{4}}\right)$$

$$= 1-e^{-\frac{3}{4}}-e^{-\frac{1}{2}}+e^{-\frac{5}{4}} = 0.2076.$$

Por lo tanto la probabilidad de que durante el primer año se presente una reclamación por choque de una mujer y en los primeros tres años se presente una reclamación por choque de un hombre es de aproximadamente el 20%.

Ejemplo 4.1.3 Esperanza matemática

Al día las cantidades de dinero reclamadas por daños a bicicletas aseguradas son independientes con fdp común

$$f_C(c) = \begin{cases} 4c^{-5} & c > 1 \\ 0 & c.o.c. \end{cases}$$

Con c la cantidad en miles de pesos. ¿Cuál es el valor esperado de la mayor cantidad reclamada en las siguientes cuatro reclamaciones?
Sean C_1, C_2, C_3 y C_4 las variables aleatorias que representan la cantidad reclamada en los siguientes cuatro reclamos respectivamente, sí se define

$$C^+ = máx(C_1, C_2, C_3, C_4)$$

Es necesario obtener $E[C^+]$, primero se calcula la fda y fdp de C^+

$$\begin{aligned} F_{C^+}(c) &= P\left(C^+ \leq c\right) \\ &= P[máx(C_1, C_2, C_3, C_4) \leq c] \\ &= P[C_1 \leq c, C_2 \leq c, C_3 \leq c, C_4 \leq c] \end{aligned}$$

por independencia de las reclamaciones

$$\begin{aligned} &= P(C_1 \leq c)P(C_2 \leq c)P(C_3 \leq c)P(C_4 \leq c) \\ &= [F_C(c)]^4 \end{aligned}$$

Con $F_C(c)$ la fda de cada C_i para $i = 1, 2, 3, 4$, se procede a obtener la fdp de C^+

$$f_{C^+}(c) = 4[F_C(c)]^3 f_C(c)$$

Como $f_C(c) = 4c^{-5}$ para $c > 1$ por el teorema fundamental del cálculo

$$F_C(c) = \int_1^c 4t^{-5}dt = -t^{-4}\Big|_1^c = 1-c^{-4}$$

$$f_{C^+}(c) = 4[F_C(c)]^3 f_C(c)$$

al sustituir $F_C(c)$ y $f_C(c)$

$$
\begin{aligned}
&= 4\left(1-c^{-4}\right)^3\left(4c^{-5}\right)\\
&= 16c^{-5}\left(1-3c^{-4}+3c^{-8}-c^{-12}\right)\\
&= 16\left(c^{-5}-3c^{-9}+3c^{-13}-c^{-17}\right)
\end{aligned}
$$

Ahora ya es posible calcular $E\left[C^+\right]$

$$
\begin{aligned}
E\left[C^+\right] &= 16\int_1^\infty c\left(c^{-5}-3c^{-9}+3c^{-13}-c^{-17}\right)dc\\
&= 16\int_1^\infty c^{-4}-3c^{-8}+3c^{-12}-c^{-16}dc\\
&= 16\left(-\frac{c^{-3}}{3}+\frac{3c^{-7}}{7}-\frac{3c^{-11}}{11}+\frac{c^{-15}}{15}\right)\Bigg|_1^\infty\\
&= 16\left(\frac{1}{3}-\frac{3}{7}+\frac{3}{11}-\frac{1}{15}\right)\\
&= 1.773
\end{aligned}
$$

Como c está dado en miles de pesos, se espera que de las próximas cuatro reclamaciones, la más costosa será por 1773 pesos.

Ejemplo 4.1.4 Esperanza matemática II

Una compañía que exporta 2 productos, nota que tiene una perdida X del primer producto y una perdida Y del segundo producto, la compañía nota que dichas perdidas se presentan bajo la siguiente distribución conjunta

$$
f_{X,Y}(x,y)=\begin{cases}\frac{1}{4}(x+2y) & 0<y<1\ \ y\ 0<x<2\\ 0 & c.o.c.\end{cases}
$$

La compañía contrata un seguro y la póliza de dicho seguro cubre las pérdidas de la empresa, dicha póliza tiene un deducible de 1.5. Calcular el valor esperado del pago que hará la aseguradora.

Sea Q la variable aleatoria que representa el pago que hará la aseguradora. Es necesario obtener $E(Q)$.

Como la aseguradora paga el total de las perdidas $X+Y$ con un deducible de 1.5, para que la aseguradora pague es preciso que $X+Y>1.5$ con lo que que la fdp de Q es

$$
f_Q(q)=\begin{cases}x+y-\frac{3}{2} & x+y>\frac{3}{2}\\ 0 & c.o.c.\end{cases}
$$

y la esperanza es de la forma

$$E(Q) = \int \int_R f_{X,Y}(x,y) f_Q(q)\, dxdy$$

Como $x+y > \frac{3}{2}$ $0 < y < 1$ y $0 < x < 2 \Rightarrow 1 > y > \frac{3}{2} - x$ y $\frac{3}{2} > x > \frac{1}{2}$ ademas $2 > x > \frac{3}{2}$ y $0 < y < 1$ (figura 4.1.1), por lo tanto

$$E(Q) = \int_{\frac{1}{2}}^{\frac{3}{2}} \int_{\frac{3}{2}-x}^{1} \frac{1}{4}(x+2y)\left(x+y-\frac{3}{2}\right) dydx + \int_{\frac{3}{2}}^{2} \int_{0}^{1} \frac{1}{4}(x+2y)\left(x+y-\frac{3}{2}\right) dydx$$

$$= \frac{1}{4} \int_{\frac{1}{2}}^{\frac{3}{2}} \int_{\frac{3}{2}-x}^{1} x^2 + 3xy - \frac{3}{2}x + 2y^2 - 3ydydx + \frac{1}{4} \int_{\frac{3}{2}}^{2} \int_{0}^{1} x^2 + 3xy - \frac{3}{2}x + 2y^2 - 3ydydx$$

$$= \frac{11}{96} + \frac{27}{96} = \frac{19}{48} = .3958$$

Entonces la aseguradora tiene una esperanza de pago de .3958.

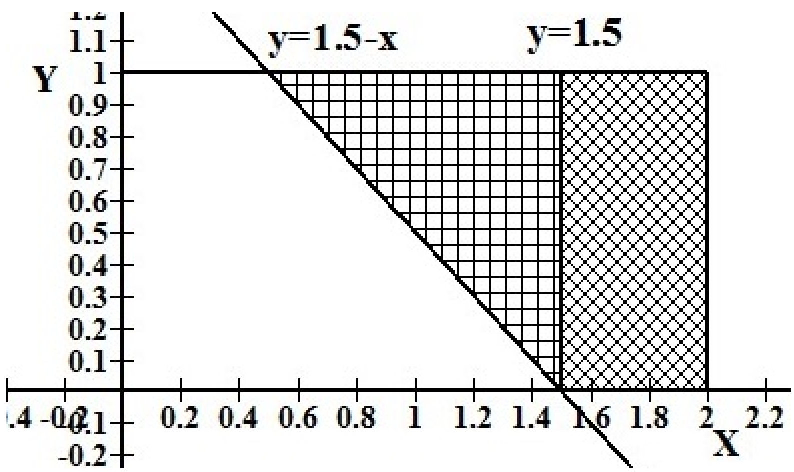

FIGURA 4.1.1. Región del ejemplo 4.1.4

Ejemplo 4.1.5 Probabilidad condicional

Una compañía aseguradora hace un estimado de pago X sobre un siniestro. Cuando sucede el siniestro y esté se valora, la compañía paga la cantidad valorada Y.
La compañía a determinado que X y Y tienen $fdp\ conjunta$

$$f_{X,Y}(x,y) = \frac{2}{x^2(x-1)} y^{\frac{-(2x-1)}{x-1}} \quad \begin{array}{l} x > 1 \\ y > 1 \end{array}.$$

a)Dado que el estimado de pago es de 3, determinar la probabilidad de que el pago final sea de entre 2 y 4.
b)Calcular $E[Y \mid X = n]$ tal que $n\ \varepsilon\ \{1, 2, \ldots, 10\}$

a)Es necesario obtener $P[2 \leq Y \leq 4 \mid X = 3]$.
Se calcula $f_X(3)$ y $f_{X,Y}(3,y)$

$$f_X(3) = \int_1^{\infty} \frac{1}{9} y^{-\frac{5}{2}} dy = \frac{2}{27}$$

$$f_{X,Y}(3,y) = \frac{1}{9} y^{-\frac{5}{2}}$$

Por último

$$\begin{aligned} P[2 \leq Y \leq 4 \mid X = 3] &= \int_2^4 \frac{f_{X,Y}(3,y)}{f_X(3)} dy \\ &= \int_2^4 \frac{\frac{1}{9} y^{-\frac{5}{2}}}{\frac{2}{27}} dy \\ &= -\left. \frac{1}{\sqrt{y^3}} \right|_2^4 \\ &= 0.2285. \end{aligned}$$

b)Se encuentra la fmp marginal de X y con ella, la función de distribución condicional $f(y \mid x)$

$$f_X(x) = \int_1^{\infty} \frac{2}{x^2(x-1)} y^{\frac{1-2x}{x-1}} dy = \frac{2}{x^3}$$

$$f(y \mid x) = \frac{f(x,y)}{f_X(x)} = \frac{\frac{2}{x^2(x-1)} y^{\frac{-(2x-1)}{x-1}}}{\frac{2}{x^3}} = \frac{x}{x-1} y^{\frac{1-2x}{x-1}}$$

Ahora se puede calcular la esperanza condicional

$$E[Y \mid X] = \int_1^{\infty} y \left(\frac{x}{x-1} y^{\frac{1-2x}{x-1}} \right) dy = \int_1^{\infty} \frac{x}{x-1} y^{\frac{-x}{x-1}} dy = x.$$

Por lo tanto $E[Y \mid X = n] = n$ tal que $n \, \varepsilon \, \{1, 2, \ldots, 10\}$ lo que quiere decir que la compañía tiene la esperanza de pagar el n estimado al momento del siniestro.

4.2. Economía

Ejemplo 4.2.1 Producto del ingreso marginal

Sea un productor que emplea m personas para producir un total de q unidades de un producto por día. Se observa que q está en función de m. Si r es el ingreso total que el fabricante recibe al vender esas unidades, entonces r también se considera como una función de m. Así, se puede ver a $\frac{dr}{dm}$ como la razón de cambio del ingreso con respecto al número de empleados. La derivada $\frac{dr}{dm}$ se llama producto del ingreso marginal, y es aproximadamente igual al cambio en el ingreso que resulta cuando un fabricante emplea un trabajador adicional.

Un fabricante determina que m empleados producirán un total de q unidades de un producto por día, donde

$$q = \frac{10m^2}{\sqrt{m^2+19}}.$$

a)Si la ecuación de demanda para el producto es $p = \frac{900}{(q+9)}$, determinar el producto del ingreso marginal cuando $m = 9$.
b)La producción de unidades tiene asociada una probabilidad de defecto del 5%. ¿Cúal es la probabilidad de que solo existan 2 unidades defectuosas con nueve empleados trabajando?

a)La cuestión es encontrar $\frac{dr}{dm}$, donde r es el ingreso. Al aplicar la regla de la cadena,

$$\frac{dr}{dm} = \frac{dr}{dq} \cdot \frac{dq}{dm}.$$

Por lo tanto se debe encontrar $\frac{dr}{dq}$ y $\frac{dq}{dm}$ cuando $m = 9$.
La función de ingreso[2]está dada por

$$\begin{aligned} r &= pq \\ &= \left(\frac{900}{q+9}\right) q \\ &= \frac{900q}{q+9} \end{aligned}$$

Al derivar, por la regla del cociente

$$\begin{aligned} \frac{dr}{dq} &= \frac{(q+9)(900) - 900q(1)}{(q+9)^2} \\ &= \frac{8100}{(q+9)^2}. \end{aligned}$$

Primero se utiliza la ecuación $q = \frac{10m^2}{\sqrt{m^2+19}}$ para encontrar el valor correspondiente de q y se evaluá en $m = 9$

$$q = \frac{10(9)^2}{\sqrt{(9)^2+19}} = 81.$$

Se tiene que,

$$\left.\frac{dr}{dq}\right|_{m=9} = \left.\frac{dr}{dq}\right|_{q=81} = \frac{8100}{(81+9)^2} = 1.$$

[2]El ingreso que resulta de una o más transacciones comerciales es el pago total recibido, y a veces se la llama ingreso bruto. Si $r(q)$ es el ingreso por vender q artículos al precio de p cada uno, entonces r es la función lineal $r(q) = pq$ y el precio de venta p también puede ser visto como la demanda para el producto.

Ahora se calcula $\frac{dq}{dm}$.

$$
\begin{aligned}
\frac{dq}{dm} &= \frac{d}{dm}\left(\frac{10m^2}{\sqrt{m^2+19}}\right) \\
&= \frac{\left(m^2+19\right)^{\frac{1}{2}} \frac{d}{dm}\left(10m^2\right) - \left(10m^2\right) \frac{d}{dm}\left[\left(m^2+19\right)^{\frac{1}{2}}\right]}{\left[\left(m^2+19\right)^{\frac{1}{2}}\right]^2} \\
&= \frac{\left(m^2+19\right)^{\frac{1}{2}}(20m) - \left(10m^2\right)\left[\frac{1}{2}\left(m^2+19\right)^{\frac{-1}{2}}(2m)\right]}{m^2+19},
\end{aligned}
$$

por lo tanto

$$
\begin{aligned}
\left.\frac{dq}{dm}\right|_{m=9} &= \frac{(81+19)^{\frac{1}{2}}(20\cdot 9) - (10\cdot 81)\left[\frac{1}{2}(81+19)^{\frac{-1}{2}}(2\cdot 9)\right]}{81+19} \\
&= 10.71
\end{aligned}
$$

Por lo tanto, si el fabricante contrata un décimo trabajador, su ingreso aumentará en aproximadamente \$10.71 por día.

b)Las unidades defectuosas siguen una distribución Poisson. Sea X la variable aleatoria que representa una unidad defectuosa, para nueve trabajadores se tiene una producción de 81 piezas, como hay una probabilidad del 5% de que haya unidades defectuosas, se esperan 4 unidades defectuosas.

$$
f_X(x) = \frac{e^{-\lambda}\lambda^x}{x!}
$$

$$
P[X=2|\lambda=4] = \frac{e^{-4}4^2}{2!} = .0732
$$

Hay una pequeña probabilidad del 7% de que solo salgan dos unidades defectuosas.

Ejemplo 4.2.2 Distribución normal

El precio de los últimos diez días de las acciones de la compañía de refrescos más importante en México se muestran en la siguiente tabla.

Día	Precio
15/03/2016	160.15
16/03/2016	162.43
17/03/2016	161.68
18/03/2016	163.36
22/03/2016	163.51
23/03/2016	163.44
28/03/2016	161.42
29/03/2016	164.85
30/03/2016	166.67
31/03/2016	166.80

Si se supone que los precios siguen una distribución $Normal$. ¿Cuál es la probabilidad de que al siguiente día el precio de las acciones se encuentre entre los \$170 y \$180?

Como los precios se distribuyen como una variable aleatoria, se tiene que su media es $\mu = \frac{1}{n}\sum_{i=1}^{10} X_i = \frac{1634.3}{10} = 163.43$ y su desviación estándar es $s = \sqrt{\frac{1}{9}\sum_{i=1}^{10}(X_i-\mu)^2} = 2.179.$

Fecha	Precio	$(X_i\text{-media})^2$
15/03/2016	160.15	10.76
16/03/2016	162.43	1.00
17/03/2016	161.68	3.07
18/03/2016	163.36	0.01
22/03/2016	163.51	0.01
23/03/2016	163.44	0.00
28/03/2016	161.42	4.04
29/03/2016	164.85	2.01
30/03/2016	166.67	10.49
31/03/2016	166.80	11.35
Total	**1,634.31**	**42.74**

Por lo tanto los precios tienes una varianza de 4.75.
Sea Y la variable aleatoria que representa el precio de las acciones, Y se distribuye como una variable aleatoria $Normal\,(163.43, 4.75)$ y su fdp es

$$f_Y(y) = \frac{1}{4.75\sqrt{2\pi}} e^{-\left(\frac{y-163.43}{4.75\sqrt{2}}\right)^2}$$

Es necesario obtener $P[170 \leq Y \leq 180]$

$$P[170 \leq Y \leq 180] = \int_{170}^{180} \frac{1}{4.75\sqrt{2}\sqrt{\pi}} e^{-\left(\frac{y-163.43}{4.75\sqrt{2}}\right)^2} dy$$

Está integral es del tipo de (1.1.4).
Sea $x = \frac{y-163.43}{4.75\sqrt{2}} \Longrightarrow \frac{dx}{dy} = \frac{dy}{4.75\sqrt{2}}$ la integral se transforma a

$$\int_{170}^{180} \frac{1}{4.752\sqrt{2}\sqrt{\pi}} e^{-\left(\frac{y-163.43}{4.75\sqrt{2}}\right)^2} dy = \frac{1}{\sqrt{\pi}} \int_{.98}^{2.47} e^{-x^2} dx$$

Sea

$$A = \int_{.98}^{2.47} e^{-x^2} dx \Rightarrow A^2 = \left(\int_{.98}^{2.47} e^{-x^2} dx\right)^2 = \left(\int_{.98}^{2.47} e^{-x^2} dx\right)\left(\int_{.98}^{2.47} e^{-y^2} dy\right)$$

$$A^2 = \left(\int_{.98}^{2.47} e^{-x^2} dx\right)\left(\int_{.98}^{2.47} e^{-y^2} dy\right) = \int_{.98}^{2.47}\int_{.98}^{2.47} e^{-\left(x^2+y^2\right)} dxdy$$

Al hacer $x = rcos\theta$ y $y = rsen\theta$ y después de que se establecen los limites de integración se tiene que

$$A^2 = \int_{1.39}^{3.49}\int_{0}^{\frac{\pi}{2}} e^{-r^2} rd\theta dr = -\frac{\pi}{4}\left(e^{-12.18} - e^{-1,93}\right) = .1139 \quad \Rightarrow A = .3376$$

Como se tiene que

$$\int_{170}^{180} \frac{1}{4.752\sqrt{2}\sqrt{\pi}} e^{-\left(\frac{y-163.43}{4.75\sqrt{2}}\right)^2} dy = \frac{1}{\sqrt{\pi}} A = \frac{.3376}{\sqrt{\pi}} = .1904$$

Entonces la probabilidad de que la día siguiente las acciones tengan un precio que oscile entre los \$170 y \$180 es tan solo del 19%.

Ejemplo 4.2.3 Distribución binomial negativa

Una compañía minera llevo a cabo un estudio que indica que en la perforación de una mina, se tiene el 30% de probabilidad de encontrar oro.
a)Cuál es la probabilidad de encontrar oro hasta la cuarta mina perforada?
b)Cuál es la probabilidad de que el tercer descubrimiento venga en la novena mina perforada?
c)Cuál es la media y la varianza de el número de minas que deben ser perforadas si la compañía quiere establecer 3 minas productoras de oro?.

Sea X la variable aleatoria que indica que la mina perforada no contenía oro. Al perforar una mina soló existen dos posibles resultados, que contenga oro o que no contenga, se llamará como éxito con probabilidad $p = 0.3$ que la mina perforada contenga oro y como fracaso con probabilidad $1 - p = 0.7$ que la mina no tenga oro. Es notorio que la variable aleatoria X se adapta a una distribución Binomial Negativa con parámetro $p = 0.3$ y $r = r - esimo\ exito$.

a) Se necesita $P[X=3]$ y se observa que en este caso $r=1$, por lo tanto

$$P[X=3]=\binom{1+3-1}{3}(0.3)^1(0.7)^3=0.1029.$$

Esto es, hay un 10% de probabilidad de que en la cuarta mina perforada se encuentre oro por primera vez.

b) En este caso, $r=3$ y como antes ya se habían dado 2 éxitos, entonces se debe obtener $P[X=6]$, con lo que

$$P[X=6]=\binom{3+6-1}{6}(0.3)^3(0.7)^6=0.08894$$

Soló hay aproximadamente el 1% de probabilidad de que el tercer hallazgo de oro se haga en la novena mina perforada.

c) Aquí $r=3$ y se pide esperanza y varianza de X

$$E[X]=\frac{3\,(0.7)}{0.3}=7, \qquad Var[X]=\frac{3\,(0.7)}{(0.3)^2}=23.3$$

Por lo tanto se tiene la esperanza de que al realizar la perforación de la séptima mina, se encuentre por tercera vez oro.

Ejemplo 4.2.4 Distribución weibull

Un hospital tiene dos plantas generadoras de luz, el tiempo en días que pasa hasta la falla de cada planta es independiente y sigue una distribución *Weibull* con media $\Gamma\left(\frac{3}{2}\right)$.
En un corte de luz, el hospital empezara a usar la segunda planta inmediatamente después de que la primer planta falle. ¿Cuál es la varianza del tiempo total que las plantas producirán electricidad?

Sea X la variable aleatoria que representa el tiempo que la planta 1 funciona y Y la variable aleatoria que representa el tiempo que la planta 2 funciona.
Es necesario calcular $Var(X+Y)$ y como X y Y son independientes

$$Var(X+Y)=Var(X)+Var(Y)$$

se sabe que X y Y se distribuyen como una variable aleatoria $Weibull(2,1)$, por lo tanto

$$Var(X)=Var(Y)=\left[\Gamma(2)-\Gamma^2\left(\frac{3}{2}\right)\right]$$

De las relaciones $\left(\Gamma\left(n+\frac{1}{2}\right)=\frac{1\cdot3\cdot5\cdot\ldots\cdot(2n-1)\sqrt{\pi}}{2^n}\right)$[3] y $\Gamma(n)=(n-1)!$

[3] N. N. Lebedev. (1965). Special functions and their applications. Prentice Hall. (p. 19)

$$= \left[\Gamma(2) - \Gamma^2\left(\frac{3}{2}\right)\right] = \left[1 - \left(\frac{\sqrt{\pi}}{2}\right)^2\right] = 1 - \frac{\pi}{4}.$$

Con lo que

$$Var(X+Y) = 1 - \frac{\pi}{4} + 1 - \frac{\pi}{4} = 2 - \frac{\pi}{2} = .4292$$

Dado que cada una de las dos plantas generadoras de luz tienen una esperanza de $\frac{\sqrt{\pi}}{2} = .8862$, la esperanza total de que las dos plantas produzcan luz es de $\sqrt{\pi} = 1.7724$ días con una varianza de .4292 días.

4.3. Meteorología

Ejemplo 4.3.1 Distribución Gumbell

La media del máximo caudal anual de agua en cierta zona es de $28.31\frac{m^3}{s}$ y la desviación estándar es de $11.32\frac{m^3}{s}$. La última inundación que se registro en está zona ocurrió hace 50 años y durante esta inundación el caudal de agua fue de $46.46\frac{m^3}{s}$.
a)¿Cuál es la probabilidad de que la zona reciba un caudal 10% menor al que recibió durante la ultima inundación?
b)¿Cuál es la probabilidad de que no haya una inundación como la de hace 50 años en los próximos 5 años?

a)Sea X la variable aleatoria que mide el caudal de agua que la zona recibe. Como sólo hay un dato de una inundación, se asume que X tiene una distribución $Gumbell(\mu, \beta)$.
Para una distribución $Gumbell(\mu, \beta)$[4]

$$E[x] = \mu + \gamma\beta \;\; y \;\; Var[x] = \frac{\pi^2}{6}\beta^2.$$

EL problema dice que

$$E[x] = 28.31\frac{m^3}{s} \;\; y \;\; Var[x] = 11.32\frac{m^3}{s}$$

Al igualar los términos y despejar μ y β

$$\frac{\pi^2}{6}\beta^2 = 1.32\frac{m^3}{s}$$

$$\mu + \gamma\beta = 28.31\frac{m^3}{s}$$

Se tiene $\mu = 23.21\frac{m^3}{s}$ y $\beta = 8.83\frac{m^3}{s}$, con esto, X se distribuye como una variable aleatoria $Gumbell\left(23.21\frac{m^3}{s}, 8.83\frac{m^3}{s}\right)$ y es necesario obtener $P[X = 38.21]$

[4]La esperanza de una variable aleatoria $Gumbell(\mu, \beta)$ es $E[x] = \mu + \gamma\beta$, con $\gamma = constante\ de\ Euler - Mascheroni \simeq 0.5772$.

$$f_X(38.21) = e^{-e^{-\left(\frac{38.21-23-21}{8.83}\right)}}$$
$$= .8328.$$

Por lo tanto la probabilidad de que la zona reciba un caudal de agua de $38.21\frac{m^3}{s}$ es del 83.28%.

b) Sea Y la variable aleatoria que representa el número de inundaciones como la de hace 50 años. Se observa que Y se distribuye como una variable aleatoria $Binomial(n, p)$.
Se tiene que $n = 5$ y $p = \frac{1}{50} = 0.02$

$$P[Y=0] = \binom{5}{0}(0.02)^0(0.98)^5$$
$$= \left(\frac{5!}{5!(0)!}\right)(1)(0.9039)$$
$$= 0.9039$$

Es muy alta la probabilidad de que no haya una inundación como la de hace 50 años en 5 años.

Ejemplo 4.3.2 Esperanza condicional

Un estudio determina que el número anual de huracanes que golpean los estados de Guerrero y Oaxaca están conjuntamente distribuidos de la siguiente manera:

		Número anual de huracanes en Oaxaca			
		0	1	2	3
Número anual de huracanes en Guerrero	0	0.12	0.06	0.05	0.02
	1	0.13	0.15	0.12	0.03
	2	0.05	0.15	0.1	0.02

Calcular la esperanza condicional y varianza condicional del número anual de huracanes que golpearán al estado de Oaxaca dado que ningún huracán golpeará el estado de Guerrero.

Sea G la variable aleatoria que representa el número de huracanes que golpean el estado de Guerrero y O la variable aleatoria que representa el número de huracanes que golpean al estado de Oaxaca.
Es preciso obtener $E[O \mid G=0]$ y $Var[O \mid G=0]$.

Primero se obtiene la fmp marginal de $G=0$

$$P_G(0) = \sum_O p(O,0) = p(O=0,0) + p(O=1,0) + p(O=2,0) + p(O=3,0)$$

$$= 0.12 + 0.06 + 0.05 + 0.02 = 0.25$$

De (2.3.3) se tiene que $E[O \mid G=0] = \sum_{O=0}^{3} O \frac{p(O,0)}{P_G(0)}$ por lo tanto

$$E[O \mid G=0] = \sum_{O=0}^{3} O \frac{p(O,0)}{0.25} = \frac{0(0.12)}{0.25} + \frac{1(0.06)}{0.25} + \frac{2(0.05)}{0.25} + \frac{3(0.02)}{0.25} = 0.88.$$

Por (2.3.4), $Var[O \mid G=0] = E\left[O^2 \mid G=0\right] - [E[O \mid G=0]]^2$se procede a encontrar $E\left[O^2 \mid G=0\right] = \sum_{O=0}^{3} O^2 \frac{p(O,0)}{P_G(0)}$

$$E\left[O^2 \mid G=0\right] = \sum_{O=0}^{3} O^2 \frac{p(O,0)}{0.25} = \frac{0(0.12)}{0.25} + \frac{1^2(0.06)}{0.25} + \frac{2^2(0.05)}{0.25} + \frac{3^2(0.02)}{0.25} = 1.76.$$

Así

$$Var[O \mid G=0] = 1.76 - (0.88)^2 = 0.99.$$

Esto indica que dado que ningún huracán golpeara el estado de Guerrero, es muy probable que 1 huracán golpee el estado de Oaxaca.

4.4. Varios

Ejemplo 4.4.1 Proceso Poisson

En la terminal Indios Verdes de la línea 3 del metro, las salidas por hora de los trenes son de acuerdo a una distribución exponencial de parámetro $\lambda_1 = 1$, mientras que en la terminal Indios Verdes del metrobus, las salidas por hora de los camiones son de acuerdo a un proceso Poisson de parámetro $\lambda_2 = 7$.
¿Cuál es la probabilidad de que exactamente 3 camiones del metrobus salgan antes que un tren del metro?

Sea T la variable aleatoria que cuenta el número de trenes que salen por hora de Indios Verdes, por lo tanto X de distribuye como una variable aleatoria $exp(\lambda_1)$ y C la variable aleatoria que cuenta el número de camiones que salen por hora de Indios Verdes, a su vez T se distribuye como una variable aleatoria $Poisson(\lambda_2)$. Se necesita obtener

$$P(3C < T)$$

Sea X la variable aleatoria que representa la probabilidad de observar n-salidas después del tiempo t. Por lo tanto X representa la probabilidad de observar la salida de 3 camiones . Por el ejemplo (3.4.2), X se distribuye como una variable aleatoria $Gamma\left(3, \frac{1}{\lambda_2}\right)$ y ahora se requiere

$$P(X<T)$$

La $fdp\ conjunta$ del vector aleatorio (X,T) es

$$f_{X,T}(x,t)=\frac{\lambda_1 e^{-\lambda_1 t}\lambda_2^3 x^2 e^{-\lambda_1 x}}{\Gamma(3)} \quad \begin{array}{c}\lambda_1,\lambda_2>0\\ x,t\geq 0\end{array}$$

Se obtiene la probabilidad sobre la $fdp\ conjunta$

$$\begin{aligned}
P(X<T) &= \int_0^\infty \int_x^\infty \frac{\lambda_1 e^{-\lambda_1 t}\lambda_2^3 x^2 e^{-\lambda_1 x}}{\Gamma(3)} dtdx \\
&= \int_0^\infty -\left.\frac{e^{-\lambda_1 t}\lambda_2^3 x^2 e^{-\lambda_1 x}}{\Gamma(3)}\right|_x^\infty dx \\
&= \int_0^\infty \frac{e^{-\lambda_1 x}\lambda_2^3 x^2 e^{-\lambda_1 x}}{\Gamma(3)} dx \\
&= \int_0^\infty \frac{\lambda_2^3 x^2 e^{-x(\lambda_2+\lambda_1)}}{\Gamma(3)} dx \\
&= \frac{\lambda_2^3}{(\lambda_2+\lambda_1)^3}\int_0^\infty \frac{(\lambda_2+\lambda_1)^3 x^2 e^{-x(\lambda_2+\lambda_1)}}{\Gamma(3)} dx
\end{aligned}$$

Se nota el hecho de que $\frac{(\lambda_2+\lambda_1)^3 x^2 e^{-x(\lambda_2+\lambda_1)}}{\Gamma(3)}$ es la fdp de una variable aleatoria $Gamma(3,(\lambda_2+\lambda_1))$ y por tanto, por ser fdp, $\int_0^\infty \frac{(\lambda_2+\lambda_1)^3 x^2 e^{-x(\lambda_2+\lambda_1)}}{\Gamma(3)} dx = 1$, con lo que

$$P(X<T)=\frac{\lambda_2^3}{(\lambda_2+\lambda_1)^3}$$

Al sustituir $\lambda_1=1$ y $\lambda_2=7$ se tiene

$$P(X<T)=\left(\frac{7}{8}\right)^3=.6699.$$

y significa que hay aproximadamente una probabilidad del 67% de que salgan tres camiones del metrobus antes que un tren del metro.

Ejemplo 4.4.2 Estadísticos de orden

Tres personas quedan de verse a las 2pm., cada persona llega al lugar de la cita de manera independiente en un tiempo aleatorio distribuido uniformemente entre las 2 y las 3 de la tarde. Cada persona espera 10 minutos y, si no llega alguna otra, se va, en cambio, si se encuentran dos de ellas, esperan a la otra hasta las 3 pm.

a) Calcular la probabilidad de que ningún par de ellas se encuentre.
b) Cuál es la probabilidad de que se encuentren las tres personas?.
c) Cuál es la probabilidad de que no se encuentren las tres personas?.

Sean $X_1,\ X_2,\ X_3$ las variables aleatorias independientes que representan la hora de llegada de cada persona de tal forma que X_i de distribuyen $Uniforme(0,60)$, $i=1,2,3$. Y sean $X_{(1)},X_{(2)},X_{(3)}$ los estadísticos de orden con $0<X_{(1)}<X_{(2)}<X_{(3)}<60$. La $fdp\ conjunta$ de $X_{(1)},X_{(2)},X_{(3)}$ será

$$f_{X_{(1)},X_{(2)},X_{(3)}}(x_1,x_2,x_3)=3!\left(\frac{1}{60}\right)^3 \qquad 0<X_{(1)}<X_{(2)}<X_{(3)}<60.$$

a) Para que ningún par de personas se encuentre es necesario que

$$\begin{aligned}P\left[X_{(3)}-X_{(2)}>10,X_{(2)}-X_{(1)}>10\right]&=P\left[X_{(3)}>X_{(2)}+10,X_{(2)}>X_{(1)}+10\right]\\&=\int_0^{40}\int_{x_1+10}^{50}\int_{x_2+10}^{60}6\left(\frac{1}{60}\right)^3dx_3dx_2dx_1\\&=\frac{8}{27}.\end{aligned}$$

b) Si las dos primeras personas llegan en un lapso de 10 minutos, esperan al tercero que forzosamente llegará antes o hasta las tres, esto es $\left(X_{(2)}-X_{(1)}<10\right)$, y si la primer persona llega entre 2.50pm. y 3pm., implica también que las tres personas se encontrarán, con lo que $\left(50<X_{(1)}<X_{(2)}<X_{(3)}<60\right)$, por lo tanto la probabilidad deseada es

$$\begin{aligned}&P\left[X_{(2)}-X_{(1)}<10\right]+P\left[50<X_{(1)}<X_{(2)}<X_{(3)}<60\right]=\\&P\left[X_{(2)}<X_{(1)}+10\right]+P\left[50<X_{(1)}<X_{(2)}<X_{(3)}<60\right]=\\&\int_0^{50}\int_{x_1}^{x_1+10}\int_{x_2}^{60}6\left(\frac{1}{60}\right)^3dx_3dx_2dx_1+\int_{50}^{60}\int_{x_1}^{60}\int_{x_2}^{60}6\left(\frac{1}{60}\right)^3dx_3dx_2dx_1=\\&\frac{90}{216}+\frac{1}{216}=\\&\frac{91}{216}.\end{aligned}$$

c) Para que no se encuentren las tres personas, es necesario que la llegada de las primeras dos personas difiera en más de 10 minutos, $\left(X_{(2)}-X_{(1)}>10\right)$, de este modo la hora de llegada de la tercer persona es irrelevante, ya que a lo sumo se podrán encontrar dos personas, por lo tanto es preciso obtener

$$\begin{aligned} P\left[X_{(2)} - X_{(1)} > 10\right] &= P\left[X_{(2)} > X_{(1)} + 10\right] \\ &= \int_0^{50} \int_{x_1+10}^{60} \int_{x_2}^{60} 6\left(\frac{1}{60}\right)^3 dx_3 dx_2 dx_1 \\ &= \frac{125}{216}. \end{aligned}$$

Por axiomas de probabilidad se tiene que $(P[A^c] = 1 - P[A])$.[5] Entonces sí la probabilidad de que se encuentren las tres personas es de $\frac{91}{216}$, la probabilidad de que no se encuentren las tres personas será $1 - \frac{91}{216} = \frac{125}{216}$, que coincide con la obtenida en el inciso c.

Ejemplo 4.4.3 Máxima verosimilitud
Los alumnos de primer semestre de la carrera de Actuaría cursan 6 materias obligatorias. Los siguientes datos muestran el número de materias aprobadas al término del primer semestre en Actuaría.

Número de materias aprobadas	*Frecuencia de alumnos*
1	169
2	57
3	14
4	7
5	2
6	4

Puede ser asumido que los datos siguen una distribución $Geometrica(p)$. Estimar y comparar el valor observado contra la frecuencia esperada.

Sea X_i la variable que cuenta el número de materias aprobadas por alumno, del ejemplo (3.6.1), el estadístico máximo verosímil para está prueba es $\hat{p} = \frac{n}{\sum_{i=1}^{n} x_i}$, en este caso se tiene $n = 253$ y $\sum_{i=1}^{n} x_i = (1)169 + (2)57 + (3)14 + 4(7) + (5)2 + (6)4 = 387$, con lo que

$$\hat{p} = \frac{253}{387} = .6537$$

La frecuencia esperada de X_i se obtiene mediante la fórmula $f_e(X_i) = n\hat{p}(1-\hat{p})^{i-1}$. Con lo que.

[5] Casella, G., Berger, R.L. (2002). Statistical inference. Pacific Grove CA: Duxbury. (p. 10)

Número de materias aprobadas	Frecuencia de alumnos	Frecuencia esperada
1	169	165.4
2	57	57.3
3	14	19.8
4	7	6.9
5	2	2.4
6	4	0.8

Por lo tanto se observa que los casos con mayor diferencia fueron los de alumnos con 3 y 6 materias aprobadas, en el primer caso fueron menos los alumnos que aprobaron y en el segundo caso que fueron más los alumnos que aprobaron con respecto a los esperados.

Conclusiones

Las matemáticas que se enseñan en las materias de cálculo I, II, III y IV junto con herramientas de lógica y álgebra son la base para poder entender las herramientas de las que la probabilidad y estadística se sirven para su estudio, mostrándole al lector la necesidad de conocer y dominar en lo mejor posible las matemáticas, ya que sin ellas la teoría de la probabilidad y estadística carecen de sentido.

De acuerdo a lo expuesto en cada capítulo, el plan de estudios de la carrera de actuaria tiene un sentido lógico, ya que el proceso de aprendizaje consiste en primero conocer las matemáticas universitarias, que no solo son aplicaciones, sino también toda su base teórica, la que generará una capacidad analítica para poder adaptar el comportamiento de los fenómenos a una función conocida o tratar de generar una nueva que se adapte al fenómeno en cuestión.

El segundo paso es adentrar al campo de estudio de la probabilidad y estadística, en donde se muestra la teoría que rige a ambas materias, los postulados con los que se construyen los modelos, sus limitantes, sus condiciones y el porqué de estos, esto genera habilidad en el manejo de los datos que se trabajan con las técnicas del cálculo para poder generar una teoría estadística y probabilística con cimientos matemáticamente probados.

Una vez dominada la teoría se puede trabajar con problemas o situaciones que se presentan en la vida diaria, tratando de buscar patrones o comportamientos que se adaptan a la teoría que ya se ah comprendido, cuando se sabe identificar el problema lo importante es darle solución de la mejor manera posible, y esto se logra gracias a que ya se conocen varias técnicas para atacar el problema.

Se observa que no todos los ejercicios teóricos tienen su respectiva aplicación, esto se debe a que aún hay mucho campo de investigación en donde el actuario puede actuar, otro aspecto es que como se señalo en la introducción, mientras mayor sea la muestra, mejores resultados se obtendrán, esto nos dice que tal vez nuevas teorías se están estudiando y están en proceso de ser establecidas.

La mayoría de las aplicaciones conllevan a la toma de decisiones que tendrán un impacto financiero, siendo este uno de los objetivos perseguidos, ver que el campo de trabajo exige conocer y dominar a fondo estos temas para obtener el resultado deseado. Es así que una base sólida en matemáticas generará más herramientas al momento de enfrentarse a modelos probabilísticos y estadísticos, que al ser analizados y resueltos, tendrán directa o indirectamente un impacto financiero.

Apéndice A

Tabla de distribuciones

Distribuciones discretas

$Binomial\,(n,p)$
$fmp \quad P(X=x)=\binom{n}{x}p^x(1-p)^{n-x}; \quad x=0,1,2,\ldots,n; \quad 0\le p\le 1$
$Esperanza \quad E[x]=np$
$Varianza \quad Var[x]=np(1-p)$

$Binomial\ negativa\,(r,p)$
$fmp \quad P(X=x)=\binom{r+x-1}{x}p^r(1-p)^x; \quad x=0,1,2,\ldots,n; \quad 0\le p\le 1$
$Esperanza \quad E[x]=\frac{r(1-p)}{p}$
$Varianza \quad Var[x]=\frac{r(1-p)}{p^2}$

$Geometrica\,(p)$
$fmp \quad P(X=x)=p(1-p)^{x-1}; \quad x=1,2,\ldots,n; \quad 0\le p\le 1$
$Esperanza \quad E[x]=\frac{1}{p}$
$Varianza \quad Var[x]=\frac{1-p}{p^2}$

$Poisson\,(\lambda)$
$fmp \quad P(X=x)=\frac{e^{-\lambda}\lambda^x}{x!}; \quad x=0,1,2,\ldots,n; \quad 0\le\lambda<\infty$
$Esperanza \quad E[x]=\lambda$
$Varianza \quad Var[x]=\lambda$

Distribuciones continuas

$Cauchy\,(\theta,\sigma)$
$fdp \quad f_X\,(x)=\frac{1}{\pi\sigma}\frac{1}{1+\left(\frac{x-\theta}{\sigma}\right)^2}; \quad -\infty<x<\infty;\ -\infty<\theta<\infty;\ \sigma>0$
$Esperanza \quad No\ \ existe$
$Varianza \quad No\ \ existe$

$Exponencial\,(\lambda)$
$fdp \quad f_X\,(x)=\lambda e^{-\lambda x}; \quad 0\leq x<\infty;\ \lambda>0$
$Esperanza \quad E\,[x]=\lambda$
$Varianza \quad Var\,[x]=\lambda^2$

$Gamma\,(\alpha,\beta)$
$fdp \quad f_X\,(x)=\frac{x^{\alpha-1}e^{-\frac{x}{\beta}}}{\Gamma(\alpha)\beta^{\alpha}}; \quad 0\leq x<\infty;\ \alpha,\beta>0$
$Esperanza \quad E\,[x]=\alpha\beta$
$Varianza \quad Var\,[x]=\alpha\beta^2$

$Normal\,\left(\mu,\sigma^2\right)$
$fdp \quad f_X\,(x)=\frac{1}{\sqrt{2\pi}\sigma}e^{\frac{-(x-\mu)^2}{2\sigma^2}}; \quad -\infty<x,\mu<\infty;\ \sigma>0$
$Esperanza \quad E\,[x]=\mu$
$Varianza \quad Var\,[x]=\sigma^2$

$Uniforme\,(a,b)$
$fdp \quad f_X\,(x)=\frac{1}{b-a}; \quad a\leq x\leq b$
$Esperanza \quad E\,[x]=\frac{b+a}{2}$
$Varianza \quad Var\,[x]=\frac{(b-a)^2}{2}$

$Weibull\,(\alpha,\beta)$
$fdp \quad f_X\,(x)=\frac{\alpha}{\beta}x^{\alpha-1}e^{\frac{-x^{\alpha}}{\beta}}; \quad 0\leq x<\infty;\ \alpha,\beta>0$
$Esperanza \quad E\,[x]=\beta^{\frac{1}{\alpha}}\Gamma\left(1+\frac{1}{\alpha}\right)$
$Varianza \quad Var\,[x]=\beta^{\frac{2}{\alpha}}\left[\Gamma\left(1+\frac{2}{\alpha}\right)-\Gamma^2\left(1+\frac{1}{\alpha}\right)\right]$

Distribuciones continuas

$Gumbell\,(\mu,\beta)$

$fdp \quad f_X\,(x) = e^{-e^{-\left(\frac{x-\mu}{\beta}\right)}}; \quad 0 \leq x < \infty;\ \mu,\beta > 0$

$Esperanza \quad E\,[x] = \mu + \gamma\beta$

$Varianza \quad Var\,[x] = \frac{\pi^2\beta^2}{6}$

$\gamma = constante\ de\ Euler - Mascheroni \simeq 0.5772.$

Referencias

Apostol, T. M. (2009). Calculus Vol. II. Barcelona: Reverte.

Bartle, Robert G., (2010). Introducción al análisis matemático de una variable. 3^a ed. México: Limusa Wiley.

Casella, G., Berger, R.L. (2002). Statistical inference. Pacific Grove CA: Duxbury.

Claudio Pita Ruiz. (1995). Cálculo Vectorial. México. Prentice Hall.

Courant, R. Fritz, J. (1999). Introducción al cálculo y al análisis matemático Vol. I. México: Limusa

Courant, R. & John, F. (2008). Introduction to calculus and analysis Vol. II. Beijing: World Books Publishing Corporation.

George B. Thomas, Jr. (2006). Cálculo. Una variable. Undécima edición. México. Pearson.

Harris, J.;Hirst, J.L.; Mossinghoff, M. (2008). Combinatorics and graph theory. Springer.

Mood, A.M.; Graybill, F.A, Boes, D.C. (1974). Introduction to the theory of statistics. Nueva York: McGraw-Hill.

N. N. Lebedev. (1965). Special functions and their applications. Prentice Hall.

Rincón. L. (2007). Curso Intermedio de Probabilidad. México. C.U. UNAM.

Ross, S. M. (2006). A first course in probability theory. New Jersey: Prentice Hall.

Spivak, M. (2003). Calculus, 2^a ed. México: Reverté

Varian, Hal R. (1992). Análisis microeconómico. 3^a ed. Barcelona. Antoni Bosch, editor.