Definición 1.1

Un vector aleatorio n-dimensional es una función $(X_1,\dots,X_n)$ desde el espacio muestral Ω en $\mathbb{R}^n$, el espacio Euclidiano n-dimensional:

$$\begin{array}{rccc}(X_1,\dots,X_n): & \Omega & \longrightarrow & \mathbb{R}^n \\ & \omega & \longrightarrow & (X_1(\omega),\dots,X_n(\omega)).\end{array}$$

Es decir, $(X_1,\dots,X_n)$ es un vector aleatorio en $(\Omega, \mathcal{A}, P)$ ssi $X_1,\dots,X_n$ son variables aleatorias en $(\Omega, \mathcal{A}, P)$.

Definición 1.2

Distribución de probabilidad conjunta La distribución de probabilidad conjunta de $X_1,\dots,X_n$ es la colección de probabilidades, $P\{(X_1,\dots,X_n)\in B\}$, para todos los subconjuntos B de $\mathbb{R}^n$.

Definición 1.3

Función de distribución conjunta La función de distribución (acumulada) conjunta de $X_1,\dots,X_n$, se define como,

$$F_{X_1,\dots,X_n}(x_1,\dots,x_n) = P(X_1\le x_1,\dots,X_n\le x_n), \quad (x_1,\dots,x_n)\in\mathbb{R}^n.$$

Note que $F_{X_1,\dots,X_n}:\mathbb{R}^n \longrightarrow [0,1]$.

Teorema 1.1

Sea $F_{X,Y}(x,y)$ la función de distribución conjunta de X e Y. Entonces,

F1a) $\lim_{x,y\to\infty} F_{X,Y}(x,y)=1$, ambos argumentos.

F1b) $\lim_{x\to-\infty} F_{X,Y}(x,y) = \lim_{y\to-\infty} F_{X,Y}(x,y)=0$, para cada valor del otro argumento.

F2) $F_{X,Y}(x,y)$ es no decreciente en cada uno de sus argumentos.

F3) $F_{X,Y}(x,y)$ es continua por la derecha en cada uno de sus argumentos.

F4) Si $a_1<b_1$ y $a_2<b_2$, entonces,

$$\underbrace{F_{X,Y}(b_1,b_2)-F_{X,Y}(a_1,b_2)-F_{X,Y}(b_1,a_2)+F_{X,Y}(a_1,a_2)}_{P_{X,Y}\{(a_1,b_1]\times(a_2,b_2]\}=P(a_1<X\le b_1,a_2<Y\le b_2)}\ge 0$$

Definición 2.1

Vector aleatorio discreto. Un vector aleatorio $(X_1,\dots,X_n)$ se dice discreto si su recorrido, $\mathcal{X}$, es un subconjunto contable (finito o infinito) de $\mathbb{R}^n$; es decir, $(X_1,\dots,X_n)$ es dicreto ssi las coordenadas $X_1,\dots,X_n$ son variables aleatorias discretas.

De este modo, se tiene que,

i) $f_{X_1,\dots,X_n}(x_1,\dots,x_n)\ge 0$ para todo $(x_1,\dots,x_n)$,

ii) $\sum\cdots\sum_{\{(x_1,\dots,x_n)\in\mathbb{R}^n\}} f_{X_1,\dots,X_n}(x_1,\dots,x_n)=1$,

iii) $P_{X_1,\dots,X_n}(B)=\sum\cdots\sum_{\{(x_1,\dots,x_n)\in B\}} f_{X_1,\dots,X_n}(x_1,\dots,x_n)$ para todo subconjunto B de $\mathbb{R}^n$.

Definición 2.2

Vector aleatorio continuo. Un vector aleatorio $(X_1,\dots,X_n)$ se dice (absolutamente) continuo si existe una función no negativa, $f_{X_1,\dots,X_n}:\mathbb{R}^n\longrightarrow\mathbb{R}$, tal que

$$F_{X_1,\dots,X_n}(x_1,\dots,x_n)=\int_{-\infty}^{x_1}\cdots\int_{-\infty}^{x_n} f_{X_1,\dots,X_n}(u_1,\dots,u_n)du_n\cdots du_1,$$

para todo $(x_1,\dots,x_n)\in\mathbb{R}^n$.

i) $f_{X_1,\dots,X_n}(x_1,\dots,x_n)\ge 0$ para todo $(x_1,\dots,x_n)$,

ii) $\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} f_{X_1,\dots,X_n}(x_1,\dots,x_n)dx_1\cdots dx_n=1$,

iii) $P_{X_1,\dots,X_n}(B)=\int\cdots\int_{\{(x_1,\dots,x_n)\in B\}} f_{X_1,\dots,X_n}(x_1,\dots,x_n)dx_n\cdots dx_1$ para cada subconjunto B de $\mathbb{R}^n$.

Definición 1.1

Sea $(X_1,\dots,X_n)$ un vector aleatorio (arbitrario) con fda conjunta $F_{X_1,\dots,X_n}$ y fda's marginales $F_{X_1},\dots,F_{X_n}$, respectivamente. Entonces, las variables aleatorias $X_1,\dots,X_n$ se dicen (mutuamente) independientes, ssi:

$$F_{X_1,\dots,X_n}(x_1,\dots,x_n)=\prod_{i=1}^n F_{X_i}(x_i) \quad \forall\,(x_1,\dots,x_n)\in\mathbb{R}^n.$$

Definición 1.2

Si $(X_1,\dots,X_n)$ es un vector aleatorio (discreto o continuo) con fmp conjunta (c.d.) o fdp conjunta (c.c.) $f_{X_1,\dots,X_n}$, y fmp's marginales (c.d.) o fdp's marginales (c.c.) $f_{X_1},\dots,f_{X_n}$, respectivamente. Entonces, las variables aleatorias $X_1,\dots,X_n$ se dicen (mutuamente) independientes, ssi:

$$f_{X_1,\dots,X_n}(x_1,\dots,x_n)=\prod_{i=1}^n f_{X_i}(x_i) \quad \forall\,(x_1,\dots,x_n)\in\mathbb{R}^n.$$

Teorema 1.1

Sea (X,Y) un vector aleatorio bivariado discreto con *fmp* conjunta $f_{X,Y}(x,y)$. Entonces las fmp marginal de X e Y, $f_X(x)=P(X=x)$ y $f_Y(y)=P(Y=y)$, están dadas por

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

Teorema 1.1

$$E\{g(X_1,\dots,X_n)\} = \begin{cases}\sum\cdots\sum_{(x_1,\dots,x_n)\in\mathbb{R}^n} g(x_1,\dots,x_n)f_{X_1,\dots,X_n}(x_1,\dots,x_n) & \text{c.d.}\\ \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} g(x_1,\dots,x_n)f_{X_1,\dots,X_n}(x_1,\dots,x_n)dx_1\cdots dx_n & \text{c.c}\end{cases}$$

(provisto que las sumatorias y las integrales convergan)

I) **Esperanza de funciones lineales**

Si $g(x_1,\dots,x_n)=\sum_{i=1}^n a_ix_i+b$, donde $a_1,\dots,a_n,b$ son constantes reales, entonces

$$E\left(\sum_{i=1}^n a_iX_i+b\right)=\sum_{i=1}^n a_iE(X_i)+b,$$

provisto que $E(|X_i|)<\infty$ para todo $i=1,\dots,n$.

II) **Función generadora de momentos (fgm) multivariada**

Si $g(x_1,\dots,x_n)=\prod_{i=1}^n e^{t_ix_i}=e^{\sum_{i=1}^n t_ix_i}$, entonces $E\{g(X_1,\dots,X_n)\}$ define la fgm conjunta de $X_1,\dots,X_n$, provisto, obviamente, que ella exista; es decir:

Definición 1.1

La fgm de conjunta de $X_1,\dots,X_n$ se difine como,

$$M_{X_1,\dots,X_n}(t_1,\dots,t_n)=E\left(e^{\sum_{i=1}^n t_iX_i}\right) = \begin{cases}\sum\cdots\sum_{(x_1,\dots,x_n)\in\mathbb{R}^n} e^{\sum_{i=1}^n t_ix_i}f_{X_1,\dots,X_n}(x_1,\dots,x_n) & \text{c.d}\\ \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} e^{\sum_{i=1}^n t_ix_i}f_{X_1,\dots,X_n}(x_1,\dots,x_n)dx_1\cdots dx_n & \text{c.c}\end{cases}$$

provisto que la esperanza exista para todo $(t_1,\dots,t_n)\in\mathbb{R}^n$ tal que $|t_k|<h_k$, algún $(h_1,\dots,h_k)\in\mathbb{R}^n$, con $h_k>0$ para todo $k=1,\dots,n$.

Propiedades de la fgm multivariada:

(i) $M_{X_1,\dots,X_k}(t_1,\dots,t_k)=M_{X_1,\dots,X_k,X_{k+1},\dots,X_n}(t_1,\dots,t_k,0,\dots,0)$ para todo $k=1,\dots,n-1$.

(ii) $\frac{\partial^{k_1+\cdots+k_n}M_{X_1,\dots,X_n}(t_1,\dots,t_n)}{\partial t_1^{k_1}\cdots\partial t_n^{k_n}}\Big|_{t_1=\cdots=t_n=0}=E(X_1^{k_1}\times\cdots\times X_n^{k_n})$.

(iii) $X_1,\dots,X_n$ son va's independientes si, y sólo si,

$$M_{X_1,\dots,X_n}(t_1,\dots,t_n)=\prod_{i=1}^n M_{X_i}(t_i)$$

para todo $(t_1,\dots,t_n)$ donde las fgm's existen.

(iv) Si $Y=\sum_{i=1}^n a_iX_i+b$, entonces,

$$M_Y(t)=e^{bt}M_{X_1,\dots,X_n}(a_1t,\dots,a_nt).$$

Definición 1.2

La convarianza entre dos variables aleatorias X e Y se define como,

$$\mathrm{Cov}(X,Y)=E\{(X-\mu_X)(Y-\mu_Y)\} = \begin{cases}\sum_{x\in\mathbb{R}}\sum_{y\in\mathbb{R}}(x-\mu_X)(y-\mu_Y)f_{X,Y}(x,y) & \text{c.d.}\\ \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x-\mu_X)(y-\mu_Y)f_{X,Y}(x,y)dydx & \text{c.c.}\end{cases}$$

Interpretación:

$$\sigma_{XY}>0 \implies \text{asociación } +,$$
$$\sigma_{XY}<0 \implies \text{asociación } -,$$
$$\sigma_{XY}=0 \implies \text{no hay asociación lineal.}$$

Propiedades de la covarianza: Para dos va's X e Y definidas sobre el mismo espacio de probabilidades, se tiene que:

(i) $\mathrm{Cov}(X,Y)=E(XY)-E(X)E(Y)$ (formula alternativa)

(ii) $\mathrm{Cov}(X,X)=\mathrm{Var}(X)\ge 0$ (operador positivo definido)

(iii) $\mathrm{Cov}(X,Y)=\mathrm{Cov}(Y,X)$ (simetría)

(iv) $\mathrm{Cov}(X,c)=\mathrm{Cov}(c,X)=0$ para cualquier constante $c\in\mathbb{R}$

(v) Si X e Y son va's independientes, entonces $\mathrm{Cov}(X,Y)=0$

Nota: Si X e Y son independientes, entonces $g(X)$ y $h(Y)$ también son va's idependientes, para cualquier funciones g y h

Definición 1.3

La correlación (coeficiente de correlación) entre dos variables aleatorias X e Y, se define como,

$$\rho_{XY}=\frac{\mathrm{Cov}(X,Y)}{\sqrt{\mathrm{Var}(X)\mathrm{Var}(Y)}}=\frac{\sigma_{XY}}{\sigma_X\sigma_Y}.$$

Teorema 1.2

Suponga que X es una variable aleatoria tal que $0<\sigma_X^2<\infty$. Si existen constantes $a\neq 0$ y b tal que $Y=aX+b$, entonces, $\rho_{XY}=1$ si $a>0$ (asociación lineal $+$) y $\rho_{XY}=-1$ si $a<0$ (asociación lineal $-$).

IV) **Varianza de funciones lineales**

Considerando nuevamente la funcion lineal $Y=\sum_{i=1}^n a_ix_i+b$, ahora podemos calcular su varianza:

Sean $X_1,\dots,X_n$ va's con varianzas finitas y $a_1,\dots,a_n,b$ constantes en $\mathbb{R}$. Vimos que

$$E\left(\sum_{i=1}^n a_iX_i+b\right)=\sum_{i=1}^n a_iE(X_i)+b. \quad (*)$$

Ahora, podemos también afirmar que,

$$\mathrm{Var}\left(\sum_{i=1}^n a_iX_i+b\right)=\sum_{i=1}^n a_i^2\mathrm{Var}(X_i)+2\sum_{\substack{1\le i,j\le n\\ i<j}} a_ia_j\mathrm{Cov}(X_i,X_j). \quad (**)$$

Definición 1.1

Esperanza de un vector aleatorio Sea $\boldsymbol{X}=(X_1,\dots,X_n)^T$ un vector aleatorio de dimensión n. La esperanza de $\boldsymbol{X}$, denotado como $E(\boldsymbol{X})$, se define como,

$$E(\boldsymbol{X})=(E(X_1),E(X_2),\cdots,E(X_n))^T.$$

Es decir, si $\boldsymbol{\mu}:=E(\boldsymbol{X})$ y $\mu_i=E(X_i)$ para $i=1,\dots,n$, entonces $\boldsymbol{\mu}=(\mu_1,\dots,\mu_n)^\top$ y también se llama **vector esperado**.

Definición 1.2

Matriz de varianza-covarianza Sea $\boldsymbol{X}=(X_1,\dots,X_n)^T$ un vector aleatorio de dimensión n. La matriz de varianza-covarianza de $\boldsymbol{X}$ se define como,

$$\mathrm{Var}(\boldsymbol{X})=E\left([\boldsymbol{X}-E(\boldsymbol{X})][\boldsymbol{X}-E(\boldsymbol{X})]^T\right) = \begin{pmatrix}\mathrm{Var}(X_1) & \mathrm{Cov}(X_1,X_2) & \dots & \mathrm{Cov}(X_1,X_n)\\ \mathrm{Cov}(X_2,X_1) & \mathrm{Var}(X_2) & \dots & \mathrm{Cov}(X_2,X_n)\\ \vdots & \vdots & \ddots & \vdots\\ \mathrm{Cov}(X_n,X_1) & \mathrm{Cov}(X_n,X_2) & \cdots & \mathrm{Var}(X_n)\end{pmatrix}.$$

Definición 1.3

Matriz de covarianza Sean $\boldsymbol{X}=(X_1,\dots,X_n)^T$ e $\boldsymbol{Y}=(Y_1,\dots,Y_m)^T$ dos vectores aleatorios definidos en un mismo espacio de probabilidad. La matriz de covarianza entre $\boldsymbol{X}$ e $\boldsymbol{Y}$ se define como,

$$\mathrm{Cov}(\boldsymbol{X},\boldsymbol{Y})=E\left([\boldsymbol{X}-E(\boldsymbol{X})][\boldsymbol{Y}-E(\boldsymbol{Y})]^T\right) = \begin{pmatrix}\mathrm{Cov}(X_1,Y_1) & \mathrm{Cov}(X_1,Y_2) & \dots & \mathrm{Cov}(X_1,Y_m)\\ \mathrm{Cov}(X_2,Y_1) & \mathrm{Cov}(X_2,Y_2) & \dots & \mathrm{Cov}(X_2,Y_m)\\ \vdots & \vdots & \ddots & \vdots\\ \mathrm{Cov}(X_n,Y_1) & \mathrm{Cov}(X_n,Y_2) & \cdots & \mathrm{Cov}(X_n,Y_m)\end{pmatrix}.$$

Esperanza y Varianza de Vectores Aleatorios

Propiedades básicas

1) $E(\boldsymbol{AX}+\boldsymbol{b})=\boldsymbol{A}E(\boldsymbol{X})+\boldsymbol{b}$ y $\mathrm{Var}(\boldsymbol{AX}+\boldsymbol{b})=\boldsymbol{A}\mathrm{Var}(\boldsymbol{X})\boldsymbol{A}^\top$

2a) $\boldsymbol{\Sigma}=\mathrm{Var}(\boldsymbol{X})$ es una matriz simétrica, $\boldsymbol{\Sigma}=((\sigma_{ij}))=((\sigma_{ji}))=\boldsymbol{\Sigma}^\top$

2b) $\boldsymbol{\Sigma}=\mathrm{Var}(\boldsymbol{X})$ es una matriz semidefinida positiva (s.d.p.), es decir, $\boldsymbol{a}^\top\boldsymbol{\Sigma a}\ge 0\ \forall\,\boldsymbol{a}\in\mathbb{R}^n$

2c) $\boldsymbol{\Sigma}=\mathrm{Var}(\boldsymbol{X})$ es una matriz definida positiva (d.p.), es decir, $\boldsymbol{a}^\top\boldsymbol{\Sigma a}\ge 0\ \ \forall\,\boldsymbol{a}\neq\boldsymbol{0}\iff X_1,\dots,X_n$ son linealmente independientes.

3) $\mathrm{Cov}(\boldsymbol{X},\boldsymbol{Y})=E(\boldsymbol{XY}^\top)-E(\boldsymbol{X})E(\boldsymbol{Y})^\top$, $\mathrm{Cov}(\boldsymbol{X},\boldsymbol{X})=\mathrm{Var}(\boldsymbol{X})$ y $\mathrm{Cov}(\boldsymbol{AX}+\boldsymbol{C},\boldsymbol{BY}+\boldsymbol{D})=\boldsymbol{A}\mathrm{Cov}(\boldsymbol{X},\boldsymbol{Y})\boldsymbol{B}^\top$.

Note que $\mathrm{Cov}(\boldsymbol{X},\boldsymbol{Y})=\mathrm{Cov}(\boldsymbol{Y},\boldsymbol{X})^\top$, de modo que $\mathrm{Var}(\boldsymbol{X}\pm\boldsymbol{Y})=\mathrm{Var}(\boldsymbol{X})+\mathrm{Var}(\boldsymbol{Y})\pm\mathrm{Cov}(\boldsymbol{X},\boldsymbol{Y})\pm\mathrm{Cov}(\boldsymbol{Y},\boldsymbol{X})$.

Definición 3.2

Se dice que un vector aleatorio $\boldsymbol{X}=(X_1,\dots,X_n)^\top$ tiene distribución normal n-variada con vector de medias $\boldsymbol{\mu}=(\mu_1,\dots,\mu_n)^\top$ y matriz de varianza-covarianza $\boldsymbol{\Sigma}=((\sigma_{ij}))_{i,j=1,\dots,n}$, lo cual se escribe como $\boldsymbol{X}\sim N_n(\boldsymbol{\mu},\boldsymbol{\Sigma})$, ssi

$$\boldsymbol{X}=\boldsymbol{AZ}+\boldsymbol{\mu},$$

donde $\boldsymbol{AA}^\top=\boldsymbol{\Sigma}$ y $\boldsymbol{Z}=(Z_1,\dots,Z_m)^\top\sim N_m(\boldsymbol{0},\boldsymbol{I}_m)$, es decir, $Z_1,\dots,Z_m\overset{iid}{\sim}N(0,1)$. En otras palabras,

$$\boldsymbol{X}\sim N_n(\boldsymbol{\mu},\boldsymbol{\Sigma})\overset{\text{def.}}{\iff} M_{\boldsymbol{X}}(\boldsymbol{t})=e^{\boldsymbol{t}^\top\boldsymbol{\mu}+\frac{1}{2}\boldsymbol{t}^\top\boldsymbol{\Sigma t}}, \quad \boldsymbol{t}\in\mathbb{R}^n.$$

Equivalentemente, si $\boldsymbol{\Sigma}>0$ (matriz definida positiva), entonces,

$$\boldsymbol{X}\sim N_n(\boldsymbol{\mu},\boldsymbol{\Sigma})\overset{\text{def.}}{\iff} f_{\boldsymbol{X}}(\boldsymbol{x})=(2\pi)^{-n/2}|\boldsymbol{\Sigma}|^{-1/2}e^{-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^\top\boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})}, \quad \boldsymbol{x}\in\mathbb{R}^n.$$

Propiedades básicas: Sea $\boldsymbol{X}\sim N_n(\boldsymbol{\mu},\boldsymbol{\Sigma})$. Entonces:

1) $\boldsymbol{BX}+\boldsymbol{b}\sim N_m(\boldsymbol{B\mu}+\boldsymbol{b},\boldsymbol{B\Sigma B}^\top)$ para cualquier matriz $\boldsymbol{B}\in\mathbb{R}^{m\times n}$ y vector $\boldsymbol{b}\in\mathbb{R}^{m\times 1}$.

Demostración: Use fgm.

En particular, si $\boldsymbol{\Sigma}>0$, entonces $\boldsymbol{\Sigma}^{-1/2}(\boldsymbol{X}-\boldsymbol{\mu})\sim N_n(\boldsymbol{0},\boldsymbol{I}_n)$, donde $\boldsymbol{B}=\boldsymbol{\Sigma}^{-1/2}=(\boldsymbol{\Sigma}^{1/2})^{-1}$ y $\boldsymbol{\Sigma}^{1/2}$ es la única raíz cuadrada simétrica de $\boldsymbol{\Sigma}$.

2) Considere la partición,

$$\boldsymbol{X}=\begin{pmatrix}\boldsymbol{X}_1\\ \boldsymbol{X}_2\end{pmatrix}\longrightarrow\begin{pmatrix}\boldsymbol{\mu}_1\\ \boldsymbol{\mu}_2\end{pmatrix}, \quad \begin{pmatrix}\boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12}\\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22}\end{pmatrix},$$

donde los $\boldsymbol{X}_j$'s y $\boldsymbol{\mu}_j$´s son vectores $n_j\times 1$, los $\boldsymbol{\Sigma}_{ij}$'s son matrices $n_i\times n_j$ y $n_1+n_2=n$. Entonces:

a) $\boldsymbol{X}_j\sim N_{n_j}(\boldsymbol{\mu}_j,\Sigma_{jj})$, $j=1,2$, donde $n_1+n_2=n$.

b) $\boldsymbol{X}_1$ y $\boldsymbol{X}_2\iff\mathrm{Cov}(\boldsymbol{X}_1,\boldsymbol{X}_2)=\boldsymbol{\Sigma}_{12}=\boldsymbol{\Sigma}_{21}^\top=\boldsymbol{0}$.

3) Si $\boldsymbol{\Sigma}>0$, entonces $(\boldsymbol{X}-\boldsymbol{\mu})^\top\boldsymbol{\Sigma}^{-1}(\boldsymbol{X}-\boldsymbol{\mu})\sim\chi^2_n$.

Aunque la fdp anterior parece complicada, la distribución normal bivariada es una de las más utilizadas. Algunas de sus muchas propiedades incluyen,

i) La distribución marginal de X es $N(\mu_X,\sigma_X^2)$

ii) La distribución marginal de Y is $N(\mu_Y,\sigma_Y^2)$

iii) $\rho_{XY}=0\iff X$ e Y son independientes

iv) Para constantes cualquiera a y b, la distribución de $aX+bY$ es $N(a\mu_X+b\mu_Y, a^2\sigma_X^2+b^2\sigma_Y^2+2ab\rho_{XY}\sigma_X\sigma_Y)$

Definición 2.1

Matriz de Correlación Sea $\boldsymbol{X}=(X_1,X_2,\cdots,X_n)^T$ un vector aleatorio n dimensional. La matriz de correlación de $\boldsymbol{X}$, denotada como R, se define como,

$$\mathbf{R}=\begin{pmatrix}1 & \rho_{X_1,X_2} & \cdots & \rho_{X_1,X_n}\\ \rho_{X_2,X_1} & 1 & \cdots & \rho_{X_2,X_n}\\ \vdots & \vdots & \ddots & \vdots\\ \rho_{X_n,X_1} & \rho_{X_n,X_2} & \cdots & 1\end{pmatrix}.$$

De forma similar se define la matriz de correlación $\mathbf{R}_{XY}=((\rho_{X_i,Y_j}))_{n\times m}$ entre dos vectores aleatorios $\boldsymbol{X}=(X_1,\dots,X_n)^\top$ e $\boldsymbol{Y}=(Y_1,\dots,Y_m)^\top$, de modo que la matriz $\mathbf{R}$ se tiene para $\boldsymbol{Y}=\boldsymbol{X}$.

Teorema 1.1

Sea (X_1,X_1) un vector aleatorio continuo con fdp conjunta $f_{(X_1,X_2)}$. Sean, $\mathcal{X}=\{(x_1,x_2): f_{(X_1,X_2)}(x_1,x_2)>0\}\subseteq\mathbb{R}^2$. Suponga que,

i) $y_1=g_1(x_1,x_2)$ e $y_2=g_2(x_1,x_2)$ definen una transformación uno a uno de $\mathcal{X}$ en $\mathcal{Y}=\{(y_1,y_2): y_1=g_1(x_1,x_2),\ y_2=g_2(x_1,x_2)$ para algún $(x_1,x_2)\in\mathcal{Y}\}\subseteq\mathbb{R}^2$.

ii) Las derivadas parciales de la trasnformación inversa $x_1=h_1(y_1,y_2)$ y $x_2=h_2(y_1,y_2)$ son continuas sobre $\mathcal{Y}$.

iii) El jacobiano de la transformación

$$J=\det\begin{pmatrix}\frac{\partial x_1}{\partial y_1} & \frac{\partial x_1}{\partial y_2}\\ \frac{\partial x_2}{\partial y_1} & \frac{\partial x_2}{\partial y_2}\end{pmatrix} \text{ es } \neq 0 \text{ para } (y_1,y_2)\in\mathcal{Y}.$$

Entonces la fdp conjunta de $Y_1=g_1(X_1,X_2)$ y $Y_2=g_2(X_1,X_2)$, está dada por,

$$f_{Y_1,Y_2}(y_1,y_2)=\begin{cases}|J|f_{(X,Y)}(h_1(y_1,y_2),h_2(y_1,y_2), & \text{si } (y_1,y_2)\in\mathcal{Y},\\ 0, & \text{eoc.}\end{cases}$$

Teorema 2.1

Sean $X_1,\dots,X_n$ variables aleatorias independientes. Para cada $i=1,\dots,n$, suponga que $g_i(x_i)$ es una función solo de x_i. Entonces, las transformaciones $Y_1=g_1(X_1),\dots,Y_n=g_n(X_n)$ también son variables aleatorias independientes.

Distribución uniforme discreta

Definición 2.1

Una variable aleatoria X tiene una distribución uniforme discreta sobre $\mathcal{X} = \{x_1, ..., x_N\}$, donde N es un entero positivo, si su fmp está dada por,

$$p(x) = \begin{cases} \frac{1}{N}, & \text{si } x \in \mathcal{X}, \\ 0, & \text{eoc.} \end{cases}$$

Notación: $X \sim UD(\{x_1, ..., x_N\})$.

Corolario 2.1

Si $\mathcal{X} = \{1, 2, ..., N\}$, entonces,

i) $E(X) = \dfrac{N+1}{2}$

ii) $Var(X) = \dfrac{N^2-1}{12}$

iii) $M_X(t) = \frac{1}{N}\sum_{x=1}^{N} e^{xt}$.

Distribución Exponencial

$$X \sim Exp(\lambda)$$

$$P[X = x] = \lambda e^{-\lambda x}$$

$$x \geq 0$$

$$E[X] = \frac{1}{\lambda} \qquad M(t) \boxed{=} \frac{\lambda}{\lambda - t}$$

$$Var(X) = \frac{1}{\lambda^2}$$

Definición 2.2

Una variable aleatoria X tiene una distribución binomial con parámetros n y p, si su fmp está dada por,

$$p(x) = \begin{cases} \binom{n}{x} p^x(1-p)^{n-x}, & \text{si } x = 0, 1, \cdots, n, \\ 0. & \text{eoc.} \end{cases}$$

donde n es un entero positivo y $0 < p < 1$. Si $n = 1$, la distribución binomial es denominada una distribución Bernoulli con parámetro p

Teorema 2.2

Sea $X \sim Bin(n, p)$, entonces,

i) $E(X) = np$

ii) $Var(X) = npq$,

iii) $M(t) = (pe^t + q)^n$, donde $q = 1 - p$.

Definición 2.3

Suponga que elegimos al azar y sin reemplazo n objetos desde un grupo de N, de los cuales sólo K poseen una característica de interés, digamos A. Luego, $N(\Omega) = \binom{N}{n}$. Sea X =número de objetos con la característica A de los n elegidos. Entonces la fmp de X está dada por,

$$p(x) = \begin{cases} \frac{\binom{K}{x}\binom{N-K}{n-x}}{\binom{N}{n}}, & x = 0, 1, ..., \min\{n, K\}, \\ 0, & \text{eoc.} \end{cases}$$

En este caso se dice que X tiene una Distribución Hipergeométrica.

Teorema 2.3

Sean $X \sim Hip(n, K, N)$, $p = \frac{K}{N}$ y $q = 1 - p$. Entonces,

i) $E(X) = np$

ii) $Var(X) = npq\,\dfrac{N-n}{N-1}$

iii) $p(x) \simeq \binom{n}{x} p^x(1-p)^{n-x}$, para N grande.

Definición 2.4

Poisson Distribution Se dice que una variable aleatoria X tiene una distribución de Poisson con parámetro $\lambda > 0$ si su fmp está dada por,

$$p(x) = \begin{cases} \frac{e^{-\lambda}\lambda^x}{x!}, & \text{si } x = 0, 1, \cdots, \\ 0, & \text{eoc.} \end{cases}$$

Teorema 2.4

Sea $X \sim P(\lambda)$, entonces,

i) $E(X) = \lambda$

ii) $Var(X) = \lambda$

iii) $M(t) = \exp\{\lambda(e^t - 1)\}$

Definición 2.5

Se dice que una variable aleatoria X tiene una distribución binomial negativa (o de Pascal) con parámetros r y p si su fmp esta dada por,

$$p(x) = \begin{cases} \binom{x-1}{r-1} p^r(1-p)^{x-r}, & \text{si } x = r, r+1, \ldots \\ 0, & \text{eoc} \end{cases}$$

Para $r = 1$, se dice que la variable aleatoria tiene una distribución geométrica con parámetro p, y su fmp está dada por,

$$p(x) = \begin{cases} p(1-p)^{x-1}, & \text{si } x = 1, 2, \ldots \\ 0, & \text{eoc} \end{cases}$$

Teorema 2.6

Sea $X \sim BN(r, p)$, entonces,

i) $E(X) = r/p$

ii) $Var(X) = r(1-p)/p^2$

iii) $M(t) = \left\{\dfrac{pe^t}{1-(1-p)e^t}\right\}^r$.

Corolario 2.2

Si $X \sim Geo(p)$, entonces,

i) $E(X) = 1/p$

ii) $Var(X) = (1-p)/p^2$

iii) $M(t) = \left\{\dfrac{pe^t}{1-(1-p)e^t}\right\}$.

Distribución uniforme

Definición 2.6

La distribución uniforme continua se define extendiendo la masa uniformemente en un intervalo (a, b). Su fdp está dada por,

$$f(x) = \begin{cases} \frac{1}{b-a} & \text{si } x \in (a, b) \\ 0 & \text{eoc} \end{cases}$$

Teorema 2.7

Si $X \sim U(a, b)$, entonces,

i) $E(X) = \dfrac{a+b}{2}$

ii) $Var(X) = \dfrac{(b-a)^2}{12}$

iii) $M(t) = \dfrac{e^{bt} - e^{at}}{t(b-a)}$, $\quad t \neq 0$.

Definición 2.7

Se dice que una variable aleatoria X tiene una distribución normal con parámetros μ (un número real) y σ (un real positivo), si su fdp está dada por,

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right\}, \ x \in \mathbb{R}.$$

Definición 2.8

Distribución normal estándar Si $Z \sim N(0, 1)$ se dice que Z tiene una distribución normal estándar. Su fdp está dada por,

$$\phi(z) = \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2}z^2\right\}, \ z \in \mathbb{R}.$$

Teorema 2.8

Sea $X \sim N(\mu, \sigma^2)$. Entonces,

i) $E(X) = \mu$

ii) $Var(X) = \sigma^2$

iii) $M(t) = \exp\{\mu t + \frac{1}{2}\sigma^2 t^2\}$.

Teorema 2.9

Sea $X \sim N(\mu, \sigma^2)$. Entonces,

i) $Y = a + bX \sim N(a + b\mu, b^2\sigma^2)$

ii) $Z = (X - \mu)/\sigma \sim N(0, 1)$.

Definición 2.9

Se dice que la variable aleatoria X tiene una distribución gama con los parámetros $\alpha > 0$ y $\lambda > 0$ si su fdp está dada por

$$f(x) = \begin{cases} \frac{\lambda}{\Gamma(\alpha)}(\lambda x)^{r-1}\exp(-\lambda x), & \text{si } x > 0, \\ 0, & \text{eoc} \end{cases}$$

donde $\Gamma(\cdot)$ es la función gamma, es decir, $\Gamma(\alpha) = \int_0^\infty t^{\alpha-1}\exp(-t)dt$, con $\Gamma(\alpha+1) = \alpha\Gamma(\alpha)$, $\Gamma(1/2) = \sqrt{\pi}$ y $\Gamma(n) = (n-1)!$, $n = 1, 2, \ldots$.

Teorema 2.10

Si $X \sim Gama(\alpha, \lambda)$, entonces,

i) $E(X) = \alpha/\lambda$

ii) $Var(X) = \alpha/\lambda^2$

iii) $M(t) = (1 - t/\lambda)^{-\alpha}$, si $t < \lambda$.

Propiedades

i) $\Gamma(1) = 1$.

ii) $\Gamma(p+1) = p\Gamma(p)$, $p > 0$.

iii) Si $p \in \mathbb{N}$, $\Gamma(p) = (p-1)!$.

iv) Si $k \in \mathbb{N}$, $\Gamma(p+k) = p(p+1)\cdots(p+k-1)\Gamma(p)$.

v) Si $a > 0$, $\int_0^{+\infty} x^{p-1}e^{-ax}dx = \dfrac{\Gamma(p)}{a^p}$.

vi) $\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$.

TRUCOS BUENARDOS

$$(a+b)^n = \sum_{k=0}^{n}\binom{n}{k} a^{n-k} b^k$$

Binomio de Newton

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

Serie de Euler

$$\frac{d}{dx}[x^n] = nx^{n-1}$$

Derivada de una potencia

$a, r \in \mathbb{R} \quad$ si $|r| < 1$.

$$\sum_{k=0}^{\infty} ar^k \boxed{=} \frac{a}{1-r}$$

Serie Geometrica

n OBJETOS

ARREGLOS DE TAMAÑO r

	SIN REEMPLAZO	CON REEMPLAZO
IMPORTA ORDEN	$\frac{n!}{(n-r)!} \Rightarrow$	$n^r \Rightarrow$
NO IMPORTA ORDEN	$\binom{n}{r} \Rightarrow$	$\binom{n+r-1}{r}$

Tecnicas de Conteo

$$\varphi_X(t) = E\left[e^{itX}\right] = \int_{-\infty}^{\infty} e^{itx} f_X(x)\,dx$$

$$\varphi'_X(0) = i\,E[X]$$

$$\varphi''_X(0) = -E[X^2].$$

$$\varphi^{(n)}_X(0) = i^n\,E[X^n]$$

Funcion Caracteristica

Binomial(n, p)

pmf $P(X = x|n, p) = \binom{n}{x} p^x (1-p)^{n-x}; \quad x = 0, 1, 2, \ldots, n; \quad 0 \le p \le 1$

mean and variance $\mathrm{E}X = np, \quad \mathrm{Var}\, X = np(1-p)$

mgf $M_X(t) = [pe^t + (1-p)]^n$

notes Related to Binomial Theorem (Theorem 3.2.2). The *multinomial* distribution (Definition 4.6.2) is a multivariate version of the binomial distribution.

Geometric(p)

pmf $P(X = x|p) = p(1-p)^{x-1}; \quad x = 1, 2, \ldots; \quad 0 \le p \le 1$

mean and variance $\mathrm{E}X = \frac{1}{p}, \quad \mathrm{Var}\, X = \frac{1-p}{p^2}$

mgf $M_X(t) = \dfrac{pe^t}{1-(1-p)e^t} \quad t < -\log(1-p)$

notes $Y = X - 1$ is negative binomial$(1, p)$. The distribution is *memoryless*: $P(X > s|X > t) = P(X > s - t)$.

Hypergeometric

pmf $P(X = x|N, M, K) = \frac{\binom{M}{x}\binom{N-M}{K-x}}{\binom{N}{K}}; \quad x = 0, 1, 2, \ldots, K;$
$M - (N - K) \le x \le M; \quad N, M, K \ge 0$

mean and variance $\mathrm{E}X = \frac{KM}{N}, \quad \mathrm{Var}\, X = \frac{KM}{N}\frac{(N-M)(N-K)}{N(N-1)}$

notes If $K \ll M$ and N, the range $x = 0, 1, 2, \ldots, K$ will be appropriate.

Negative binomial(r, p)

pmf $P(X = x|r, p) = \binom{r+x-1}{x} p^r (1-p)^x; \quad x = 0, 1, \ldots; \quad 0 \le p \le 1$

mean and variance $\mathrm{E}X = \frac{r(1-p)}{p}, \quad \mathrm{Var}\, X = \frac{r(1-p)}{p^2}$

mgf $M_X(t) = \left(\frac{p}{1-(1-p)e^t}\right)^r, \quad t < -\log(1-p)$

notes An alternate form of the pmf is given by $P(Y = y|r, p) = \binom{y-1}{r-1} p^r (1-p)^{y-r}$, $y = r, r+1, \ldots$. The random variable $Y = X + r$. The negative binomial can be derived as a gamma mixture of Poissons. (See Exercise 4.34.)

Poisson(λ)

pmf $P(X = x|\lambda) = \frac{e^{-\lambda}\lambda^x}{x!}; \quad x = 0, 1, \ldots; \quad 0 \le \lambda < \infty$

mean and variance $\mathrm{E}X = \lambda, \quad \mathrm{Var}\, X = \lambda$

mgf $M_X(t) = e^{\lambda(e^t - 1)}$

Exponential(β)

pdf $f(x|\beta) = \frac{1}{\beta} e^{-x/\beta}, \quad 0 \le x < \infty, \quad \beta > 0$

mean and variance $\mathrm{E}X = \beta, \quad \mathrm{Var}\, X = \beta^2$

mgf $M_X(t) = \frac{1}{1-\beta t}, \quad t < \frac{1}{\beta}$

notes Special case of the gamma distribution. Has the *memoryless* property. Has many special cases: $Y = X^{1/\gamma}$ is *Weibull*, $Y = \sqrt{2X/\beta}$ is *Rayleigh*, $Y = \alpha - \gamma \log(X/\beta)$ is *Gumbel*.

Gamma(α, β)

pdf $f(x|\alpha, \beta) = \frac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha-1} e^{-x/\beta}, \quad 0 \le x < \infty, \quad \alpha, \beta > 0$

mean and variance $\mathrm{E}X = \alpha\beta, \quad \mathrm{Var}\, X = \alpha\beta^2$

mgf $M_X(t) = \left(\frac{1}{1-\beta t}\right)^\alpha, \quad t < \frac{1}{\beta}$

notes Some special cases are exponential ($\alpha = 1$) and chi squared ($\alpha = p/2$, $\beta = 2$). If $\alpha = \frac{3}{2}$, $Y = \sqrt{X/\beta}$ is *Maxwell*. $Y = 1/X$ has the *inverted gamma distribution*. Can also be related to the Poisson (Example 3.2.1).

Normal(μ, σ^2)

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

mean and variance $\mathrm{E}X = \mu, \quad \mathrm{Var}\, X = \sigma^2$

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

notes Sometimes called the *Gaussian* distribution.

Uniform(a, b)

pdf $f(x|a, b) = \frac{1}{b-a}, \quad a \le x \le b$

mean and variance $\mathrm{E}X = \frac{b+a}{2}, \quad \mathrm{Var}\, X = \frac{(b-a)^2}{12}$

mgf $M_X(t) = \frac{e^{bt} - e^{at}}{(b-a)t}$

notes If $a = 0$ and $b = 1$, this is a special case of the beta ($\alpha = \beta = 1$).

TRUCOS BUENARDOS

$$(a + b)^n = \sum_{k=0}^{n} \binom{n}{k} a^{n-k} b^k$$

Binomio de Newton

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

Serie de Euler

$$\frac{d}{dx}[x^n] = nx^{n-1}$$

Derivada de una potencia

$a, r \in \mathbb{R}$ si $|r| < 1$

$$\sum_{k=0}^{\infty} ar^k = \frac{a}{1-r}$$

Serie Geometrica

n OBJETOS

ARREGLOS DE TAMAÑO r

	SIN REEMPLAZO	CON REEMPLAZO
IMPORTA ORDEN	$\frac{n!}{(n-r)!} \Rightarrow$	$n^r \Rightarrow$
NO IMPORTA ORDEN	$\binom{n}{r} \Rightarrow$	$\binom{n+r-1}{r}$

Tecnicas de Conteo

$$\varphi_X(t) = \mathrm{E}\left[e^{itX}\right] = \int_{-\infty}^{+\infty} e^{itx} f_X(x)\, dx$$

$$\varphi'_X(0) = i\,\mathrm{E}[X]$$

$$\varphi''_X(0) = -\mathrm{E}[X^2]$$

$$\varphi^{(n)}_X(0) = i^n\,\mathrm{E}[X^n]$$

Funcion Caracteristica

Propiedades

i) $\Gamma(1) = 1$.

ii) $\Gamma(p+1) = p\Gamma(p)$, $p > 0$.

iii) Si $p \in \mathbb{N}$, $\Gamma(p) = (p-1)!$.

iv) Si $k \in \mathbb{N}$, $\Gamma(p+k) = p(p+1)\cdots(p+k-1)\Gamma(p)$.

v) Si $a > 0$, $\int_0^{+\infty} x^{p-1} e^{-ax}\, dx = \frac{\Gamma(p)}{a^p}$.

vi) $\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$.

Funcion Gamma (Propiedades)

Teorema 2.9

Sea $X \sim N(\mu, \sigma^2)$. Entonces,

i) $Y = a + bX \sim N(a + b\mu, b^2\sigma^2)$

ii) $Z = (X - \mu)/\sigma \sim N(0, 1)$.

Funcion Normal (Teorema)

Teorema 1.1

Una función $f_X(x)$, $x \in \mathbb{R}$, es la fmp (o la fdp) de una variable aleatoria X si y sólo si:

a) $f_X(x) \geq 0$ para todo x ($f_X(x) > 0$ si $x \in \mathcal{X}$), y

b) $\sum_{x \in \mathcal{X}} f_X(x) = 1$ (fmp); o $\int_{-\infty}^{\infty} f_X(x)dx = 1$ (fdp).

Además, para cualquier conjunto $B \in \mathcal{B}(\mathbb{R})$, se tiene que

$$P_X(B) = P(X \in B) = \begin{cases} \sum_{x \in B} f_X(x), & \text{en el caso discreto,} \\ \int_B f_X(x)dx, & \text{en el caso continuo.} \end{cases}$$

ii) Se dice que la variable aleatoria X tiene (o sigue) una **distribución** (absolutamente) **continua**, o que X es una **variable aleatoria** (absolutamente) **continua**, si existe una función no negativa $f_X(x)$, $x \in \mathbb{R}$, llamada **función de densidad de probabilidad (fdp)**, tal que $F_X(x) = P(X \leq x) = \int_{-\infty}^{x} f_X(x)dx$ para todo $x \in \mathbb{R}$.

En este caso, el recorrido $\mathcal{X}$ de X es un conjunto no contable de números reales; $P(X = x) = F_X(x) - F_X(x^-) = 0$ para todo x; y $f_X(x) = \dfrac{dF_X(x)}{dx}$ en todos aquellos puntos x donde $F_X(x)$ es diferenciable.

Definición 2.1

La **esperanza (o valor esperado o media)** de la variable aleatoria X, se define como

$$E(X) = \begin{cases} \sum_{x \in \mathcal{X}} x f_X(x) & \text{si } X \text{ es discreta ,} \\ \int_{-\infty}^{\infty} x f_X(x)dx & \text{si } X \text{ es continua ,} \end{cases}$$

provisto que la suma o la integral existan. Si la suma o la integral divergen, o no estan definidas, se dice que $E(X)$ no existe.

Teorema 2.1

Sean a, b y c constantes, y sean X e Y variables aleatorias cuyas esperanza existen. Entonces,

E1) Si $X(\omega) = c$ (constante) para todo ω, es decir, la variable aleatori X es degenera en c ($P(X = c) = 1$), entonces $E(X) = c$

E2) Si $X(\omega) \geq 0$ para todo ω, es decir, la variable aleatoria X es no negativa, entonces $E(X) \geq 0$.

Además, de la formula general de esperanza dada en (*), sigue qu

$$E(X) = \int_0^{\infty} (1 - F_X(x))dx = \int_0^{\infty} P(X > x)dx.$$

En particular, si X es una variable aleatoria discreta con recorrido $\mathcal{X} = \mathbb{Z}_+$ (enteros no negativos), entonces

$$E(X) = \sum_{x=0}^{\infty} P(X > x) = \sum_{x=1}^{\infty} P(X \geq x). \quad (**)$$

E3) Si $X(\omega) \geq Y(\omega)$ para todo ω, entonces $E(X) \geq E(Y)$

E4) Si $a \leq X(\omega) \leq b$ para todo ω, entonces $a \leq E(X) \leq b$.

E5) $E(aX + b) = a\,E(X) + b$ (linealidad del operador $E(\cdot)$).

E6) (Aditividad) Sean X es una variable aleatoria en $(\Omega, \mathcal{A}, P)$, y sean $g_1(x), g_2(x)$, $x \in \mathbb{R}$, funciones tales que las variables aleatorias $g_1(X)$ y $g_2(X)$ tienen esperanza finita. Entonces,

$$E\{a\, g_1(X) + b\, g_2(X) + c\} = a\, E\{g_1(X)\} + b\, E\{g_1(X)\} + c,$$

donde a, b, c son constantes reales.

E7) (Desigualdad de Jensen) Sea X una variable aleatoria con esperanza finita, y g una función de $\mathbb{R}$ para $\mathbb{R}$:
i) Si g es convexa, entonces $E\{g(X)\} \geq g(E(X))$.
ii) Si g es cóncava, entonces $E\{g(X)\} \leq g(E(X))$.

Definición 2.3

Sea X una variable aleatoria y sea $\mu_X = E(X)$. La varianza de X se define como, $\mathrm{Var}(X) = E\{(X - \mu_X)^2\}$, es decir,

$$\mathrm{Var}(X) = \begin{cases} \int_{-\infty}^{\infty} (x - \mu_X)^2 f_X(x)dx & \text{si } X \text{ es continua} \\ \sum_{x \in \mathcal{X}} (x - \mu_X)^2 f_X(x) & \text{si } X \text{ es discreta .} \end{cases}$$

La raíz cuadrada positiva de $\mathrm{Var}(X)$ es la desviación estándar de X.

Teorema 2.2

Sea X una variable aleatoria cuya varianza existe, y sean a, b, c constantes reales. Entonces,

V1) $\mathrm{Var}(X) \geq 0$ y $\mathrm{Var}(X) = 0$ si y sólo si $X \equiv c$

V2) $\mathrm{Var}(X) = E(X^2) - \{E(X)\}^2$

V3) $\mathrm{Var}(aX + b) = a^2\, \mathrm{Var}(X)$

Recordemos que si X es una variable aleatoria en $(\Omega, \mathcal{A}, P)$ y g es una función con dominio y recorrido en los reales, entonces $g(X)$ también es una variable aleatoria en $(\Omega, \mathcal{A}, P)$, y su esperanza puede ser calculada como,

$$E\{g(X)\} = \begin{cases} \sum_{x \in \mathcal{X}} g(x) f_X(x) & \text{si } X \text{ es discreta ,} \\ \int_{-\infty}^{\infty} g(x) f_X(x)dx & \text{si } X \text{ es continua ,} \end{cases}$$

siempre que exista la $\sum$ o la $\int$. Como fue dicho antes, si $E(|g(X)|) = \infty$, decimos que $E\{g(X)\}$ no existe.

Definición 2.2

La **mediana** de una variable aleatoria X es algún número $m = \mathrm{med}(X)$ tal que

$$P(X \geq m) \geq \frac{1}{2} \quad \text{y} \quad P(X \leq m) \geq \frac{1}{2}.$$

Notas:

a) Si X es una variable aleatoria con media finita y distribución simétrica, entonces $\mu_X = E(X)$ es una mediana

b) Si X tiene distribución asimétrica (o simétrica de colas pesadas), la mediana de X puede ser una "mejor" medida localización.

Definición 1.1

Sea X una variable aleatoria en $(\Omega, \mathcal{A}, P)$, y asuma que existen las sumatorias o integrales requeridas para cada entero positivo k:

i) El k-ésimo momento no centrado de X, se define como,

$$E(X^k) = \begin{cases} \sum_{x \in \mathcal{X}} x^k f_X(x) & \text{si } X \text{es discreta,} \\ \int_{-\infty}^{\infty} x^k f_X(x)dx & \text{si } X \text{es continua.} \end{cases}$$

$k = 1 \Longrightarrow E(X) = \mu_X$ es la media de X.

ii) El k-ésimo momento centrado de X, se define como,

$$E\{(X - \mu_X)^k\} = \begin{cases} \sum_{x \in \mathcal{X}} (x - \mu_X)^k f_X(x) & \text{si } X \text{es discreta,} \\ \int_{-\infty}^{\infty} (x - \mu_X)^k f_X(x)dx & \text{si } X \text{es continua.} \end{cases}$$

$k = 2 \Longrightarrow E\{(X - \mu_X)^2\} = \sigma_X^2$ es la varianza de X.

Definición 1.3

Función generadora de momentos: Sea X una variable aleatoria con fda F_X. La función generadora de momento (fgm) de X (o F_X), denotada por $M_X(t)$, se define como,

$$M_X(t) = E(e^{tX}),$$

siempre que exista la esperanza para t en alguna vecindad de 0. Es decir, existe $h > 0$ tal que, para todo t en $(-h,\ h)$, $E(e^{tX})$ exista.

Si la esperanza no existe en una vecindad de 0, decimos que la fgm de X no existe.

Teorema 1.3

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

$$E(X^k) = M_X^{(k)}(0),$$

donde,

$$M_X^{(k)}(0) = \frac{d^k}{dt^k} M_X(t)\,|_{t=0}$$

Es decir, el k-ésimo momento de X es igual a la k-ésima derivada de $M_X(t)$ evaluada en $t = 0$.

Teorema 1.4

Sean $F_X(x)$ y $F_Y(y)$ dos fda cuyos momentos existen.

i) Si X e Y tienen soporte acotado, entonces $F_X(u) = F_Y(u)$ para todo u si y sólo si $E(X^r) = E(Y^r)$ para todo $r = 0, 1, 2, \ldots$.

ii) Si la fgm existe y $M_X(t) = M_Y(t)$ para todo t en alguna vecindad de 0, entonces $F_X(u) = F_Y(u)$ para todo u.

Teorema 1.5

Para constantes cualquiera a y b, la fgm de la variable aleatoria $aX + b$ está dada por,

$$M_{aX+b}(t) = e^{bt} M_X(at).$$

Ejemplo 1.1

Sea X una variable aleatoria con fdp (triangular) dada por,

$$f_X(x) = \begin{cases} 1 - |x|, & \text{si } |x| < 1, \\ 0, & \text{eoc.} \end{cases}$$

Como se dijo antes, ya que $\mathcal{X} = (-1, 1)$, entonces $E(X^k)$ existe para todo $k = 1, 2, \ldots$. Luego, $M_X(t)$ existe para todo t, y esta dada por

$$\begin{aligned} M_X(t) &= \int_{-1}^{1} e^{xt}(1 - |x|)dx \\ &= \int_{-1}^{0} e^{xt}(1 + x)dx + \int_{-0}^{1} e^{xt}(1 - x)dx \\ &= \frac{1}{t^2}\left(e^t + e^{-t} - 2\right), \quad \text{(que no estaría definida en } t = 0) \end{aligned}$$

Más precisamente, al escribir $y = g(x)$, la función $g(x)$ establece un mapeo del recorrido, $\mathcal{X}$, de la variable aleatoria X, a un subconjunto $\mathcal{Y}$ de $\mathbb{R}$, el cual define el recorrido de la variable aleatoria $Y = g(X)$. Es decir,

$$g : \mathcal{X} \to \mathcal{Y}.$$

Asociado con g, se tiene un mapeo inverso, denotado por g^{-1}, que es un mapeo de subconjuntos de $\mathcal{Y}$ a subconjuntos de $\mathcal{X}$, y está definido por,

$$g^{-1}(B) = \{x \in \mathcal{X} : g(x) \in B\}, \tag{2.1}$$

para cada subconjunto B de $\mathcal{Y}$.

Si la variable aleatoria X es discreta, es decir, su recorrido, $\mathcal{X}$, es un subconjunto contable (finito o infinito) de $\mathbb{R}$, entonces, el recorrido la variable aleatoria transformada $Y = g(X)$, es decir,

$$\mathcal{Y} = \{y : y = g(x), x \in \mathcal{X}\},$$

es también un subconjunto contable (finito on infinito) de $\mathbb{R}$; es decir, Y es también una variable aleatoria discreta. Luego, para determinar su distribución de probabilidad, es suficiente encontrar su fmp.

Más precisamente, si X es una variable aleatoria discreta, entonces $Y = g(X)$ es también una variable aleatoria discreta, con fmp dada por,

$$\begin{aligned} f_Y(y) &= P(Y = y) \\ &= P(g(X) = y) \\ &= P(X \in g^{-1}(y)) \\ &= P(\{x \in \mathcal{X} : g(x) = y\}) \\ &= \sum_{\{x \in \mathcal{X} : g(x) = y\}} P(X = x) \\ &= \begin{cases} \sum_{\{x \in \mathcal{X} : g(x) = y\}} f_X(x), & \text{si } y \in \mathcal{Y}, \\ 0, & \text{eoc.} \end{cases} \end{aligned}$$

O sea, para encontrar la fmp de Y, basta con identificar el conjunto $g^{-1}(y) = \{x \in \mathcal{X} : g(x) = y\}$, para cada $y \in \mathcal{Y}$, y luego sumar las probabilidades correspondientes.

Ejemplo 2.1

Sea X una variable aleatoria discreta con fmp dada por,

x	-1	0	1	2	3
$f_X(x)$	$\frac{1}{7}$	$\frac{2}{7}$	$\frac{1}{7}$	$\frac{2}{7}$	$\frac{1}{7}$

Sea $Y = X^2$. Aquí, $y = g(x) = x^2$, $\mathcal{X} = \{-1, 0, 1, 2, 3\}$, y la variable aleatoria discreta Y toma valores en $\mathcal{Y} = \{0, 1, 4, 9\}$. Luego, la fmp de Y es,

y	0	1	4	9
$f_Y(y)$	$\frac{2}{7}$	$\frac{2}{7}$	$\frac{2}{7}$	$\frac{1}{7}$

Por ejemplo, $P_Y(Y = 1) = P_X(X = -1) + P_X(X = 1) = \frac{1}{7} + \frac{1}{7} = \frac{2}{7}$.

Suponga, ahora, que tanto X como $Y = g(X)$ son variables aleatorias continuas. En muchos casos, es posible encontrar expresiones simples para la fda de Y en términos de la fda o la fdp de X y la función g. De hecho, la fda de $Y = g(X)$ esta dada por,

$$\begin{aligned} F_Y(y) &= P(Y \leq y) \\ &= P(g(X) \leq y) \\ &= P(\{x \in \mathcal{X} : g(x) \leq y\}) \\ &= \int_{\{x \in \mathcal{X} : g(x) \leq y\}} f_X(x)dx. \end{aligned}$$

Aunque en algunos casos resulta difícil identificar la región,

$$g^{-1}((-\infty, y]) = \{x \in \mathcal{X} : g(x) \leq y\},$$

y resolver la integral de $f_X(x)$ bajo esa región, este método es muy útil para encontrar la fdp de la variable aleatoria Y.

En particular, si g es monótona en $\mathcal{X}$, entonces,

$$\{x \in \mathcal{X} : g(x) \leq y\} = \{x \in \mathcal{X} : x \leq g^{-1}(y)\}, \quad \text{si } g \text{ es creciente,}$$
$$\{x \in \mathcal{X} : g(x) \leq y\} = \{x \in \mathcal{X} : x \geq g^{-1}(y)\}, \quad \text{si } g \text{ es decreciente.}$$

Es decir, para cada $y \in \mathcal{Y}$, la fda de $Y = g(X)$ queda definida como,

$$F_Y(y) = \begin{cases} F_X(g^{-1}(y)), & \text{si } g \text{ es creciente,} \\ 1 - F_X(g^{-1}(y)), & \text{si } g \text{ es decreciente.} \end{cases}$$

Como Y también es una variable aleatoria continua, entonces, derivando esta última expresión mediante la regla de la cadena, su fdp esta dada por,

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

Esto prueba el siguiente resultado, al notar que $\frac{d}{dy} g^{-1}(y)$ es positiva si g es creciente y negativa si g es decreciente.

Teorema 2.1

Sean X e $Y = g(X)$ variables aleatorias continuas, tales que X tiene fdp $f_X(x)$ y g una función monótona. Sean $\mathcal{X}$ e $\mathcal{Y}$ definidos como en (2.3). Suponga que $f_X(x)$ es continua sobre $\mathcal{X}$ y que $g^{-1}(y)$ tiene una derivada continua sobre $\mathcal{Y}$. Entonces, la fdp de Y está dada por,

$$f_Y(y) = \begin{cases} f_X\left(g^{-1}(y)\right) \left| \dfrac{d}{dy} g^{-1}(y) \right|, & \text{para } y \in \mathcal{Y}, \\ 0, & \text{en otros casos.} \end{cases} \tag{2.4}$$

Teorema 2.2

Sean X e $Y = g(X)$ variables aleatorias continuas, donde X tiene una fdp $f_X(x)$ con recorrido $\mathcal{X}$ definido como en (2.3). Suponga que existe una partición, $A_0, A_1, \ldots, A_k$, de $\mathcal{X}$ tal que $P(X \in A_0) = 0$ y $f_X(x)$ es continua sobre cada A_i. Además, suponga que existen funciones $g_1(x), \ldots, g_k(x)$, definidas sobre $A_1, \ldots, A_k$, respectivamente, tales que:

i) $g(x) = g_i(x)$, para $x \in A_i$;

ii) $g_i(x)$ es monótona sobre A_i;

iii) el conjunto $\mathcal{Y} = \{y : y = g_i(x)$ para algún $x \in A_i\}$ es el mismo para cada $i = 1, \ldots, k$; y

iv) $g_i^{-1}(y)$ tiene una derivada continua sobre $\mathcal{Y}$, para cada $i = 1, \ldots, k$.

Entonces,

$$f_Y(y) = \begin{cases} \sum_{i=1}^{k} f_X\left(g_i^{-1}(y)\right) \left| \dfrac{d}{dy} g_i^{-1}(y) \right|, & \text{para } y \in \mathcal{Y} \\ 0, & \text{en otros casos.} \end{cases} \tag{2.5}$$

Note que $\mathcal{X}$ puede ser dividido en conjuntos $A_1, \ldots, A_k$ tales que $g(x)$

eorema 2.3

ea X una variable aleatoria con fda $F_X(x)$ continua, y defina la ariable aleatoria Y como $Y = F_X(X)$. Entonces, $Y \sim U(0, 1)$, es decir, tiene una distribución uniforme sobre $(0,\ 1)$, de modo que $(Y \leq y) = y$ para $0 < y < 1$.

Sea $X : (\Omega, \mathcal{A}, P) \longrightarrow (\mathbb{R}, \mathcal{B}, P_X)$ una variable aleatoria mixta, con función de distribución

$$F_X(x) = \sum_{\substack{x' \in D_X \\ x' \leq x}} P(X = x') + \int_{-\infty}^{x} h_X(t)dt, \quad \forall x \in \mathbb{R}.$$

i) Existe la esperanza matemática de X si $\sum_{x \in D_X} |x| P(X = x) + \int_{\mathbb{R}} |x| h_X(x)dx < +\infty$.

ii) En caso de existir la esperanza, se define como

$$E[X] = \sum_{x \in E_X} x P(X = x) + \int_{\mathbb{R}} x h_X(x)dx.$$

Notemos que la función de distribución de una variable mixta se expresa como suma de dos funciones:

- $F_X^1(x) = \sum_{\substack{x' \in D_X \\ x' \leq x}} P(X = x')$

- $F_X^2(x) = \int_{-\infty}^{x} h_X(t)dt$