1 Momentos y funciones generadoras

Definición 1 (Media) *Sea X una variable aleatoria. La media de X o valor esperado de X, denotado por μ_X o E[X] es definida por:*
i) E[X]=$\sum_{k=1}^{\infty} x_k p_k$ si X es una variable discreta con función de masa de probabilidad p_k=P([X=x_k]) y $\sum_{k=1}^{\infty} \mid x_k \mid p_k < \infty$.
ii) E[X]=$\int_{-\infty}^{\infty} x f_X(x)dx$ si X es una variable continua con función densidad de probabilidad f_X y $\int_{-\infty}^{\infty} \mid x \mid f_X(x)dx < \infty$.
iii) E[X]=$\int_0^{\infty}[1 - F_X(x)]\ dx$ - $\int_{-\infty}^{0} F_X(x)dx$ para una variable aleatoria arbitraria con $\int_0^{\infty}[1 - F_X(x)]dx < \infty$ y $\int_{-\infty}^{0} F_X(x)dx < \infty$.
Nota; E[X] es el centro de gravedad de la masa unitaria que es determinada por la función de densidad de X.
Así la media de X es una medida de donde los valores de la variable X se centran.

Definición 1 (Media) *Sea (Ω,S,P) un espacio probabilístico (espacio de medida).*
a) Si A$\in$ S, se lama integral I_A respecto a P, y se denota $\int_\Omega I_A dP$, el número real P(A).
b) Sea X=$\sum_{i=1}^{n} a_i I_{Ai}$ una función simple donde los A_i son sucesos dos a dos disjuntos y los a_i son números reales no negativos. Se define la integral de X respecto a P, y se denota por $\int_\Omega X dP$, como el número real $\sum_{i=1}^{n} a_i P(A_i)$.
c) Si X : $\Omega \longrightarrow \bar{\mathbb{R}}$ es Borel-medible y no negativa, se define la integral de X respecto a P por $\int_\Omega X dP$=sup$\{\int_\Omega s\ dP : s$ simple y $0 \leq s \leq X\}$.
d) Si X : $\Omega \longrightarrow \bar{\mathbb{R}}$ es Borel-medible, se dice que X es P-integrable si las integrales de sus partes positivas y negativas son finitas; se dice que existe la integral de X respecto a P si al menos una de ellas es finita. En ambos casos se define $\int_\Omega X dP = \int_\Omega X^+ dP - \int_\Omega X^- dP$.
e) Si X es Borel-medible y A$\in$S, se detona $\int_A X dP = \int_\Omega (X I_A) dP$, si esta última integral existe.
Notación: $\int X(\omega)dP(\omega), \int X(\omega)P(d\omega)$ y $\int_\Omega X(\omega)dP(\omega)$ significa lo mismo que $\int_\Omega X dP$. Como P es una medida de probabilidad, entonces $\int_\Omega X dP$ se suele denotar por la forma $E_P(X)$ (o simplemente, E(X) o EX, si no hay confusión) y se llama esperanza o media de X.
Si X : $\mathbb{R} \longrightarrow \mathbb{R}$ es una variable aleatoria, su integral respecto a la medida de Lebesgue (si existe) se denota usualmente por $\int X(x)dx$ o $\int X$, o $\int_{\mathbb{R}} X(x)dx$ o también $\int_{-\infty}^{\infty} X(x)dx$.

Teorema 1 *Dado X(w)=I_A(w) para algún A$\in$S. Entonces E[X]=P(A).*

Teorema 2 *Si X es una variable aleatoria y h(X)=|X| se tiene: E[X] existe si y sólo si E[|X|] existe.*

Definición 2 *Una X es una variable simétrica sobre un punto α si P([X $\geq \alpha$ + x]) = P([X $\leq \alpha$ - x]) para todo x $\in \mathbb{R}$.*

<u>Nota</u>: Una función distribución F de X es simétrica (o una variable aleatoria X es simétrica) con α como centro de simetría si
F(α - x) = 1 - F(α + x) + P([X = α +x]) para todo $x \in \mathbb{R}$.
En particular, si X es una variable aleatoria continua y simetrica con centro en α si y sólo si satisface f(α - x) = f(α +x) para todo $x \in \mathbb{R}$.
Si α = 0, X es simetrica (o F es simétrica).

Teorema 3 *Si X es una variable aleatoria simétrica con α como centro de simetría y E[|X|] $< \infty$. Entonces E[X]=α.*

Teorema 4 *Si a y b son constantes y X es una variable aleatoria con E[|X|] $< \infty$. Entonces E[|aX+b|] $< \infty$ y E[aX+b] = aE[X]+b.*
<u>Nota</u>: E[X-μ]=0.

Teorema 5 *Si X es una variable aleatoria acotada, es decir, P([| X |$\leq$ M])=1, 0< M $< \infty$. Entonces E[X] existe.*

Teorema 6 *Si X es una variable aleatoria con P([X $\geq$ 0])=1 y E[X] existe. Entonces E[X] $\geq$ 0.*

Teorema 7 *Sea X una variable aleatoria y g una función de Borel-medible sobre $\mathbb{R}$. Dado Y=g(X). Si X es una variable aleatoria discreta, entonces*

$$E[Y] = \sum_{j=1}^{\infty} g(x_j)P([X = x_j]) \tag{1}$$

en el sentido que si cualquiera lado de (1) existe, también existe el otro lado, y los dos son iguales. Si X es una función continua con función de densidad de probabilidad f, entonces:

$$E[Y] = \int_{-\infty}^{\infty} g(x)f(x)dx \tag{2}$$

en el sentido que si cualquier de las dos integrales convergen absolutamente en (2) , también converge el otro lado, y los dos son iguales.
<u>Nota</u>: Sea X una variable aleatoria discreta. Entonces el teorema 7 dice que:

$$\textstyle\sum_{j=1}^{\infty} g(x_j)P([X = x_j]) = \sum_{k=1}^{\infty} y_k P([Y = y_k])$$

en el sentido que si cualquiera de las dos series converge absolutamente, también converge la otra y las dos sumas son iguales. Si X es una variable continua con función densidad de probabilidad f. Sea h(y) la función densidad de probabildiad de Y=g(X). Entonces segun el teorema 7,

$$\int_{-\infty}^{\infty} g(x)f(x)dx = \int_{-\infty}^{\infty} yh(y)dy$$

dado que E[$\mid g(X) \mid$] $< \infty$.
Se puede definir para otros tipos de variables aleatorias. E[g(X)] es definida como la integral Stieltjes $\int_{-\infty}^{\infty} g(x)dF_X(x)$ (dado que esta integral existe), donde $F_X(x)$ es la función de distribución acumulada de X.

Teorema 8 *Sea X una variable aleatoria (discreta o continua), y g y h funciones de Borel-medible sobre $\mathbb{R}$. Entonces:*
i) E[cg(X)] = cE[g(X)] para c constante.
ii) E[cg(X) + dh(x)] = cE[g(X)] + dE[h(X)] para c y d constantes.
iii) E[g(X)] $\leq$ E[h(X)] si g(x) $\leq$ h(x) para todo x.

Definición 3 (Varianza) *Sea X una variable aleatoria y sea μ_X=E[X] existe. La varianza de X, denotado por σ_X^2 o VAR[X], es defindo por:*
i) VAR[X]=$\sum_{k=1}^{\infty}(x_k - \mu_X)^2 p_k$ si X es una variable discreta con función de masa de probabilidad p_k=P([X=x_k]).
ii) VAR[X]=$\int_{-\infty}^{\infty}(x - \mu_X)^2 f_X(x)dx$ si X es una variable continua con función densidad de probabilidad f_X.
iii) VAR[X]=$\int_0^{\infty} 2x[1 - F_X(x) + F_X(-x)]\ dx - \mu_X^2$ para una variable aleatoria arbitraria.
Nota: La VAR[X] es una medida de la dispersión.

Definición 4 (Desviación estándar) *Si X es una variable aleatoria, la desviación estándar de X, denotada por σ_X es definida como $+\sqrt{VAR[X]}$.*

Teorema 9 *Si X es una variable aleatoria y E[X^2] existe, entonces VAR[X] = E[X^2] - (E[X]$)^2$.*

Definición 5 *Una variable aleatoria X es degenerada en el punto k si*

$$P([X=x]) = \begin{cases} 1 \quad , & x = k \\ 0 \quad , & cualquier\, otro\, lugar \end{cases}$$

Teorema 10 *Si X es una variable aleatoria discreta y E[X^2] existe. Entonces VAR(X)=0 si y sólo si X es degenerada.*

Teorema 11 *Sea X una variable aleatoria y E[X^2] existe. Entonces VAR(X) $<$ E[(X-c)^2] para cualquier c $\neq$ E[X].*

Teorema 12 *Sea X una variable aleatoria y E[X^2] existe. Entonces, VAR[aX+b] = a^2 VAR[X] para cualquier a y b constantes.*

Definición 6 *Una variable aleatoria Z es estandarizada si E[Z]=0 y VAR[Z]=1.*
<u>Nota</u>: Sea una variable aleatoria X y E[$\mid X \mid^2$] $< \infty$. Se define Z=$\frac{X-E[X]}{\sqrt{VAR[X]}} = \frac{X-\mu}{\sigma}$.
Se tiene que E[Z]=0 y VAR[Z]=1.

Definición 7 (Momentos) *Si X es una variable aleatoria, el r-ésimo momento de X (sobre el origen) o el momento de orden r, denotado por m_r, es definido como m_r=E[X^r] si el valor esperado existe.*

Definición 8 (Momentos centrales) *Si X es una variable aleatoria, el r-ésimo momento central de X sobre a es definido como E[(X-a)^r]. Si a=μ, se tiene el r-ésimo momento central sobre μ (o la media), denotado por μ_r, el cual es μ_r=E[(X-μ)^r].*
<u>Nota</u>: Si se conocen $m_1, m_2, ..., m_k$, se pueden calcular $\mu_1, \mu_2, ..., \mu_k$, y viceversa.
Se tiene $\mu_k = E[(X-\mu)^k] = m_k - \binom{k}{1}\mu m_{k-1} + \binom{k}{2}\mu^2 m_{k-2} - ... + (-1)^k \mu^k$ *y*
$m_k = E[(X-\mu+\mu)^k] = \mu_k + \binom{k}{1}\mu\mu_{k-1} + \binom{k}{2}\mu^2\mu_{k-2} + ... + \mu^k$

Definición 9 (Momentos factoriales) *Si X es una variable aleatoria, el r-ésimo momento factorial de X es definido como (r es un entero positivo): E[X(X-1)...(X-r+1)].*
<u>Nota</u>: Para algunas variables aleatorias de valores enteros discretas cuya función de masa de probabilidad contiene factoriales o coeficientes binomiales, puede ser más conveniente calcular momentos factoriales.

Definición 10 (Cuantil) *Un número x satisfaciendo P([X $\leq$ x]) $\geq$ p, P([X $\geq$ x]) $\geq$ 1 - p. 0 $<$ p $<$ 1, es llamado un cuantil de orden p (or 100p-ésimo percentil) para la variable aleatoria X (o para la función distribución F de X) y se denota por ξ_p(X).*
<u>Nota</u>: Si x es un cuantil de orden p para la variable aleatoria X con función distribución F, entonces
p $\leq$ F(x) $\leq$ p +P([X=x]).
Si X es una variable aleatoria continua, un cuantil de orden p es una solución de la ecuación F(x)=p ya que P([X=x])=0.
Si F es estrictamente creciente, la solución es única. De otra forma, pueden haber muchas (inclusive incontables) soluciones de F(x)=p, cada uno de los cuales es llamado un cuantil de orden p.

Definición 11 (Mediana) *Sea una variable aleatoria con función distribución F. Un número que satisface* $\frac{1}{2} \leq F(x) \leq \frac{1}{2}$ *+ P([X=x]) o equivalentemente P([X* $\leq$ *x])* $\geq \frac{1}{2}$ *y P([X* $\geq$ *x])* $\geq \frac{1}{2}$ *es llamado la mediana de X (o F).*

Teorema 13 *Si el momento de orden t existe para una variable aleatoria X, momentos de orden 0 < s < t existen.*

Teorema 14 *Sea X una variable aleatoria no negativa con función distribución F. Entonces, E[X] =* $\int_0^\infty$ *(1-F(x))dx, en el sentido que si cualquier lado existe, también el otro existe y los dos son iguales.*

Definición 12 *Sea X una variable aleatoria definida sobre (Ω,S,P). La función*

$$M(s) = E[e^{sX}] \tag{3}$$

es conocida como la función generadora de momentos de la variable aleatoria X si el valor esperado del lado derecho de (3) existe en alguna vecindad del origen.
Nota: M_X denota la función generadora de momentos de la variable aleatoria X.

Teorema 15 *Sean X y Y dos variables aleatorias con funciones de densidad f_X y f_Y respectivamente.*
a) Si $M_X(s)$ y $M_Y(s)$ existen y son iguales para todo s en el intervalo -h < s < h para algún h > 0. Entonces, las dos funciones de distribución F_X y F_Y son iguales.
b) Si $M_X(s)$ existe para todo s en el intervalo -h < s < h para algún h > 0 y las dos funciones de distribución F_X y F_Y son iguales, entonces $M_X(s)$ y $M_Y(s)$ son iguales para todo s en el intervalo -h < s < h para algún h > 0.

Teorema 16 *Si la función generadora de momentos M(s) de una variable aleatoria X existe para s en (-s_0,s_0) con $s_0 > 0$, entonces las derivadas de todos los órdenes existen en s=0 y pueden ser evaluadas bajo el signo de la integral, es decir,*
$M^{(k)}(s) \mid_{s=0} = E[X^k]$ *para k=1,2,...*
Nota: M(s) = $E[e^{sX}] = E[\sum_{i=0}^{\infty} \frac{1}{i!}(sX)^i] = \sum_{i=0}^{\infty} \frac{1}{i!} m_i s^i$

Teorema 17 *Sea h(X) una función no negativa Borel medible de una variable aleatoria X. Si E[h(X)] existe, entonces para cualquier* $\epsilon > 0$, $P([h(X) \geq \epsilon]) \leq \frac{E[h(X)]}{\epsilon}$.

Teorema 18 *Sea X una variable aleatoria, dado h(X)=*$\mid X \mid^r$ *y* $\epsilon = K^r$*, donde r > 0 y K > 0. Entonces* $P([\mid X \mid \geq K]) \leq \frac{E[\mid X \mid^r]}{K^r}$*,la cual es la desigualdad de Markov.*

En particular si h(X)=(X-μ)^2 y $\epsilon = K^2\sigma^2$, se tiene la desigualdad de Chebychev-Bienayme:
$P([\mid X-\mu \mid \geq K\sigma]) \leq \frac{1}{K^2}$, *donde E[X] = μ y VAR(X) = σ^2.*

Definición 13 (Función Convexa) *Una función $g : \mathbb{R} \rightarrow \mathbb{R}$ es convexa si para cualquier x,y $\in \mathbb{R}$ y cualquier t $\in$ [0,1], g(tx + (1-t)y) $\leq$ tg(x) + (1-t)g(y).*
Nota: Geométricamente significa que en cada intervalor [x,y] la gráfica de g queda por debajo del segmento que va de (x,g(x)) a (y,g(y)).

Teorema 19 (Desigualdad de Jensen) *Sea X una variable aleatoria con media E[X], y sea g una función convexa y continua. Entonces E[g(X)] $\geq$ g(E[X]).*

2 Bibliografía

- Rohatgi, V. y Saleh, A. (2001), An Introduction to Probability and Statistics. John Wiley and Sons, Inc.