1

Variables aleatorias y distribución de probabilidad

1.1. El concepto de variables aleatorias

Definición 1.1 *Sea S un espacio muestral sobre el que se encuentra definida una función de probabilidad. Sea X una función de valor real definida sobre S, de manera que transforme los resultados de S en puntos sobre la recta de los reales. Se dice entonces que X es un variable aleatoria.*

Definición 1.2 *Se dice que una variable aleatoria X es discreta si el número de valores que puede tomar es contable (ya sea finito o infinito), y si éstos pueden arreglarse en una secuencia que corresponde con los enteros positivos.*

Definición 1.3 *Se dice que una variable aleatoria X es continua si sus valores consisten en uno o más intervalos de la recta de los reales.*

1.2. Distribuciones de probabilidad de variables aleatorias discretas

Definición 1.4 *Sea X una variable aleatoria discreta. Se llamará a $p(x) = P(X = x)$ función de probabilidad de la variable aleatoria X, si satisface las siguientes propiedades:*

1. $p(x) \geq 0$ *para todos los valores x de X;*

2. $\sum_x p(x) = 1$.

Definición 1.5 *La función de distribución acumulativa de la variable aleatoria X es la probabilidad de que X sea menor o igual a un valor específico de x y está dada por:*

$$F(x) = P(X \leq x) = \sum_{x_i \leq x} p(x_i)$$

En general, la función de distribución acumulativa $F(x)$ de una variable aleatoria discreta es una función no decreciente de los valores de X, de tal manera que:

1. $0 \leq F(x) \leq 1$ para cualquier x;

2. $F(x_i) \geq F(x_j)$ si $x_i \geq x_j$;

3. $P(X > x) = 1 - F(x)$.

4. $P(X = x) = F(x) - F(x-1)$;

5. $P(x_i \geq X \geq x_j) = F(x_j) - F(x_i - 1)$

1.3. Distribuciones de probabilidad de variables aleatorias continuas

Definición 1.6 *1.* $f(x) \geq 0$, $-\infty < x < \infty$,

2. $\int_{-\infty}^{\infty} f(x)\, dx$ *y*

3. $P(a \leq X \leq b) = \int_a^b f(x)\, dx$

Para cualquier a y b, entonces $f(x)$ es la función de densidad de probabilidad de la variable aleatoria continua X.

Para la función de distribución acumulativa $F(x)$ se tiene:

$$P(X \leq x) = F(x) = \int_{-\infty}^{x} f(t)\, dt$$

Dado que para cualquier varible aleatoria continua X,

$$P(X = x) = \inf_x^x f(t)\, dt = 0, \qquad \Longrightarrow \qquad P(X \leq x) = P(X < x) = F(x)$$

La distribución acumulativa $F(x)$ es una función lisa no decreciente de los valores de la v.a. con las siguiente propiedades:

1. $F(-\infty) = 0$;

2. $F(\infty) = 1$;

3. $P(a < X < b) = F(b) - F(a)$

4. $dF(x)/dx = f(x)$.

1.4. Valor esperado de una variable aleatoria

Definición 1.7 *El valor esperado de una variable aleatoria X es el promedio o valor medio de X y está dado por:*

$$\begin{aligned} E(X) &= \sum_x xp(x) && \text{Si } x \text{ es discreta, o} \\ E(X) &= \int_{-\infty}^{\infty} xf(x)\,dx && \text{Si } x \text{ es continua.} \end{aligned}$$

En donde $p(x)$ y $f(x)$ son las funciones de probabilidad y de densidad de probabilidad, respectivamente.

En general, el valor esperado de una función $g(x)$ de la variables aleatoria X, está dado por:

$$\begin{aligned} E(g(X)) &= \sum_x g(x)p(x) && \text{Si } x \text{ es discreta, o} \\ E(g(X)) &= \int_{-\infty}^{\infty} g(x)xf(x)\,dx && \text{Si } x \text{ es continua.} \end{aligned}$$

1.4.1. Propiedades

1. El valor esperado de una constante c es el valor de la constante.

$$E(c) = \int_{-\infty}^{\infty} cf(x)\,dx = c\int_{-\infty}^{\infty} f(x)\,dx = c$$

2. El valor esperado de la cantidad $aX + b$, en donde a y b son constantes, es el producto de a por el valor esperado de x más b.

$$E(aX + b) = \int_{-\infty}^{\infty} (ax + b)f(x)\,dx = a\int_{-\infty}^{\infty} xf(x)\,dx + b\int_{-\infty}^{\infty} f(x)\,dx = aE(X) + b$$

3. El valor esperado de la suma de dos funciones $g(X)$ y $h(X)$ de X es la suma de los valores esperados de $g(X)$ y $h(X)$.

$$E\left[g(X) + h(X)\right] = \int_{-\infty}^{\infty} \left[g(x) + h(x)\right]\,dx \int_{-\infty}^{\infty} g(x)f(x)\,dx + \int_{-\infty}^{i} nftyh(x)f(x)\,dx = E\left[g(X)\right] + E\left[h(X)\right]$$

1.5. Momentos de una variable aleatoria

Los momentos de una variable aleatoria X son los valores esperados de ciertas funciones de X.

Definición 1.8 *Sea X una variable aleatoria. El r-ésimo momento de X alrededor de cero se define por:*

$$\begin{aligned} \mu_r' &= E(X^r) = \sum_x x^r p(x) && \text{Si } x \text{ es discreta, o} \\ \mu_r' &= E(X^r) = \int_{-\infty}^{\infty} x^r f(x)\,dx && \text{Si } x \text{ es continua.} \end{aligned}$$

El primer momento al rededor del cero es la media o valor esperado de la variable aleatoria. y se denota por μ; de ésta manera se tiene que $\mu_1' = \mu = E(X)$.

Definición 1.9 *Sea X una variable aleatoria. El r-ésimo momento central de X o el r-ésimo momento alrededor de la media de X se define por:*

$$\begin{aligned}\mu_r &= E(X - u)^r = \textstyle\sum_x (x-\mu)^r p(x) && \textit{Si } x \textit{ es discreta, o}\\ \mu_r &= E(X - u)^r = \int_{-\infty}^{\infty} (x-\mu)^r f(x)\, dx && \textit{Si } x \textit{ es continua.}\end{aligned}$$

El momento central de cero de cualquier variable aleatoria es uno, dado que:

$$\mu_0 = E(X-\mu)^0 = E(1) = 1$$

De manera similar, el primer momento central de cualquier variables aleatoria es cero, dado que:

$$\mu_1 = E(E-\mu)^0 = E(X) - \mu = 0$$

El segundo momento central:

$$\mu_2 = E(X-\mu)^2$$

Recibe el nombre de varianza de la variable aleatoria. Puesto que:

$$\begin{aligned}\mu_2 = Var(X) &= E\left[(X-\mu)^2\right]\\ &= E(X^2 - 2X\mu + \mu^2)\\ &= E(X^2) - 2\mu^2 + \mu^2\\ &= \mu_2^{'} - \mu^2\end{aligned}$$

La varianza de cualquier variable aleatoria es el segundo momento alrededor del origen menos el cuadrado de la media. Generalmente se denota por σ^2

Es útil notar que la varianza de una variable aleatoria X es invariable; es decir, $Var(X+b) = Var(X)$ para cualquier constante b. De manera más general, se demostrará que $Var(aX+b) = a^2 Var(X)$ para cualquiera dos constantes a y b. Por definición,

$$\begin{aligned}Var(aX+b) &= E(aX+b)^2 - E^2(aX+b)\\ &= E(a^2X^2 - 2abX + b^2) - [aE(X)+b]^2\\ &= a^2E(X)^2 + 2abE(X) + b^2 - a^2E^2(X) - 2abE(X) - b^2\\ &= a^2E(X)^2 - a^2E^2(X)\\ &= a^2\left[E(X)^2 - E^2(X)\right]\\ &= a^2Var(X)\end{aligned}$$

Una medida que compara la dispersión relativa de dos distribuciones de probabilidad es el coeficiente de variación, que está definido por:

$$V = \frac{\sigma}{\mu}$$

Expresa la magnitud de la dispersión de una variable aleatoria con respecto a su valor esperado.

El tercer momento central

$$\mu_3 = E(X-\mu)^3$$

esta relacionado con la asimetría de la distribución de probabilidad de X.

Cualquier momento central de una variable aleatoria X puede expresarse en términos de los momentos de ésta, alrededor de cero. Por definición:

$$u_r = E(X-\mu)^r$$

pero la expansión de $(X-\mu)^r$ puede expresarse como:

$$(X-\mu)^r = \sum_{i=0}^{r}(-1)^i \frac{r!}{(r-i)!i!}\mu^i x^{r-i} = \sum_{i=0}^{r}(-1)^i \frac{r!}{(r-i)!i!}\mu^i E(X^{r-i}) = \sum_{i=0}^{r}(-1)^i \frac{r!}{(r-i)!i!}\mu^i \mu'_{r-i}$$

En particular,

$$\mu_3 = \mu'_3 - 3\mu\mu'_2 + 2\mu^3$$

Para las distribuciones de probabilidad que presentan un sólo pico, si $\mu_3 < 0$, se dice que la distribución es asimétrica negativamente, si $\mu_3 > 0$, la distribución es asimétrica positivamente y si $\mu_3 = 0$, la distribución recibe el nombre de simétrica.
Una medida más apropiada de la asimetría, es el tercer momento estandarizado, dado por:

$$\alpha_3 = \frac{\mu_3}{(\mu_2)^{3/2}}$$

Que recibe el nombre de coeficiente de asimetría. Una distribución de probabilidad es asimétrica positiva, negativa o simétrica si $\alpha_3 > 0$, $\alpha_3 < 0$ o $\alpha_3 = 0$ respectivamente.

El cuarto momento central,

$$\mu_4 = E(X-\mu)^4 = \mu'_4 - 4\mu\mu'_3 + 6\mu^2\mu'_2 - 3\mu^4$$

Es una medida de qué tan puntiaguda es la distribución de probabilidad y recibe el nombre de curtosis. Al igual que para el tercer momento, es preferible emplear el cuarto momento estandarizado,

$$\alpha_4 = \frac{\mu_4}{\mu_2^2}$$

Si $\alpha_4 > 3$, la distribución de probabilidad presenta un pico relativamente alto y recibe el nombre de letocúrtica, si $\alpha_4 < 3$, la distribución es relativamente plana y recibe el nombre de platicúrtica y si $\alpha_4 = 3$, la distribución no presenta un pico muy alto i muy bajo y recibe el nombre de mesocúrtica.

En este momento se considerará el concepto de variable aleatoria estandarizada. Sea X cualquier variable aleatoria con media μ y desviación estándar σ la cantidad

$$Y = \frac{(x-\mu)}{\sigma}$$

define una variable aleatoria Y con media cero y desviación estándar uno. Esta variable recibe el nombre de **variable aleatoria estandarizada** correspondiente a X

El valor esperado de Y es cero, puesto que:

$$E\left(\frac{X-\mu}{\sigma}\right) = \frac{1}{\sigma}E(X-\mu) = 0$$

De hecho, puesto que $E(Y) = 0$, el r-ésimo momento central de Y es:

$$\begin{aligned}
\mu_r(Y) &= E(Y^r) \\
&= E\left(\frac{X-\mu}{\sigma}\right)^r \\
&= \frac{1}{\sigma^r}E(X-\mu)^r \\
&= \frac{\mu_r(X)}{\sigma^r}
\end{aligned}$$

De esta manera se tiene que:

$$\mu_r(Y) = \frac{\mu_r(X)}{[\mu_2(X)]^{r/2}}$$

Ya que

$$\sigma^r = \sqrt{[\mu_2(X)]^r}$$

es decir,

$$(\sigma^r)^2 = [\mu_2(X)]^r \implies (\sigma^2)^r = [\mu_2(X)]^r \implies (\sigma^2)^r = (\sigma^2)^r$$

de donde se tiene que $Var(Y) = \mu_2(Y) = 1$. En particular, nótese que $\alpha_3(Y) = \alpha_3(X)$ y $\alpha_4(Y) = \alpha_4(X)$. La estandarización de una variable aleatoria afecta a la media y a la varianza, pero no a los factores de forma.

1.6. Otras medidas de tendencia central y dispersión

Definición 1.10 *Para cualquier variable aleatoria X, se define a la mediana $x_{0.5}$ de X, para ser:*

$$\begin{aligned}
&P(X < x_{0.5}) \leq 1/2 \quad y \quad P(X \leq x_{0.5}) \geq 1/2 && \text{si } X \text{ es discreta, o} \\
&P(X \leq x_{0.5}) = 1/2 && \text{si } X \text{ es continua}
\end{aligned}$$

Definición 1.11 *Para cualquier variable aleatoria X, se define la moda como el valor x_m de X que maximiza la función de probabilidad si X es discreta o la función de densidad si X es continua.*

Definición 1.12 *Para cualquier variable aleatoria X, el valor cuantil X_q, de orden q. $0 < q < 1$, es el valor de X tal que:*

$$\begin{aligned}
&P(X < x_q) \leq q \quad y \quad P(X \leq x_q) \geq q && \text{si } X \text{ es discreta, o} \\
&P(X \leq x_q) = q && \text{si } X \text{ es continua}
\end{aligned}$$

Definición 1.13 *La desviación media de una variable aleatoria X es el valor esperado de la diferencia absoluta entre X y su media, y está dado por:*

$$\begin{aligned}
E|X-\mu| &= \sum |x-\mu| p(x) && \text{si } X \text{ es discreta, o} \\
E|X-\mu| &= \int_{-x}^{x} |x-\mu|\, dx && \text{si } X \text{ es continua}
\end{aligned}$$

1.7. Funciones generadoras de momentos

Definición 1.14 *Sea X una variable aleatoria. El valor esperado de $exp(tX)$ recibe el nombre de función generadora de momentos, y se denota por $m_X(t)$, si el valor esperado existe para cualquier valor de t en algún intervalo $-c < t < c$ en donde c es un número positivo. En otras palabras:*

$$m_X(t) = E[exp(tX)] = \sum_x e^{tx} p(x) \qquad \text{si } X \text{ es discreta, o}$$

$$m_X(t) = E[exp(tX)] = \int_{-\infty}^{\infty} e^{tx} f(x)\, dx \quad \text{si } X \text{ es continua}$$

Si la función generadora de momentos existe, puede demostrarse que es única y que determina por completo la distribución de probabilidad de X. En otras palabras **si dos variables aleatorias tienen la misma función generadora de momentos, entonces tienen la misma distribución de probabilidad.**

Si la función generadora de momentos existe para $-c < t < c$ entonces existen las derivadas de ésta de todas las órdenes para $t = 0$. Lo anterior asegura que $m_X(t)$ generará todos los momentos de X al rededor del origen. Para demostrar lo anterior se diferencia $m_X(t)$ con respecto a t, y se evalúa la derivada en $t = 0$. Suponiendo que pueden intercambiarse los símbolos de diferenciación y esperanza, se tiene:

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

Al tomar la segunda derivada y evaluar en $t = 0$.

$$\begin{aligned}
\left.\frac{d^2 m_X(t)}{dt^2}\right|_{t=0} &= \left.\frac{d^2}{dt^2} E[e^{tX}]\right|_{t=0} \\
&= E\left[\frac{d^2}{dt^2}(e^{tX})\right] \\
&= E\left[\frac{d}{dt}(X \cdot e^{tX})\right] \\
&= \left.E[X^2 e^{tX}]\right|_{t=0} \\
&= E(X^2) \\
&= \mu_2^{'}
\end{aligned}$$

Al continuar este proceso de diferenciación se puede deducir que se obtiene el

$$\begin{aligned}
\frac{d^r m_X(t)}{dt^r} &= \frac{d^r}{dt^r} E[e^{tX}]\Big|_{t=0} \\
&= E\left[\frac{d^r}{dt^r}(e^{tX})\right] \\
&= E[X^r e^{tX}]\Big|_{t=0} \\
&= E(X^r) \\
&= \mu_r^{'}
\end{aligned}$$

mismo resultado si se reemplaza la función exponencial por su expansión en serie de potencias.

$$E[e^{tX}] = E\left(1 + tX\frac{t^2X^2}{2!} + \ldots + \frac{t^rX^r}{r!} + \ldots\right)$$

Definición 1.15 *Sea X una variable aleatoria. El valor esperado de $exp[t(X-\mu)]$ recibe el nombre de función generadora de momentos central y denota por $m_{X-\mu}(t)$, si el valor esperado existe para cualquier t en algún intervalo $-c < t < c$ en donde c es un número positivo.*

$$m_{X-\mu}(t) = Eexp[t(X-\mu)] = \sum_x exp[t(x-\mu)]p(x) \qquad \textit{si X es discreta, o}$$

$$m_{X-\mu}(t) = Eexp[t(X-\mu)] = \int_{-\infty}^{\infty} exp[t(x-\mu)]f(x)\,dx \quad \textit{si X es continua}$$

1.8. Ejercicios

Los ejercicios se encuentra en el archivo $Ej\_Cap3.Rmd$