Repaso de Probabilidades

I) Variable Aleatoria

Una variable aleatoria es una función que asocia un suceso con un número que puede ser discreto o continuo. Si el conjunto de valores que toma la v.a. es discreto, entonces se dice que la v.a. es discreta, en otro caso se dirá continua. Cada posible valor de la variable aleatoria tiene asociada una probabilidad de ocurrencia.

a) Discreta

Distribución de probabilidades de una v.a. X:

$$P(X = x_i) = p(x_i)$$

$$\sum_W p(x_i) = 1$$

$$W : \text{Dominio de definición de X}$$

b) Continua

Función de densidad de probabilidades (f.d.p.) de una v.a. X:

$$f(x) \ ; \ x \in W$$

$$\int_W f(x)dx = 1 \quad \text{y} \quad f(x) \geq 0 \quad \forall x \in W$$

Función de distribución acumulada (F.d.a.):

$$F(x) = P(X \leq x)$$

Si *F(x)* es diferenciable: $f(x) = \frac{d}{dx}F(x)$, excepto quizá en un número finito de valores de X.

II) Esperanza

La esperanza es un promedio ponderado de los valores de una v.a., donde los "pesos" corresponden a las probabilidades asociadas a dichos valores.

a) Para una v.a. discreta

$$E(X) = \sum_W x_i p(x_i)$$

b) Para una v.a. continua

$$E(X) = \int_W x f(x) dx$$

c) Función de una v.a.

Sea $Y = H(X)$, luego:

$$E(Y) = \begin{cases} \sum_{W} H(x_i)p(x_i) & \text{si X es discreta} \\ \int_{W} H(x)f(x)dx & \text{si X es continua} \end{cases}$$

d) Propiedades

Sean *X, Y* dos v.a , *a* y *b* constantes:

i) Linealidad

$$E(aX + b) = aE(x) + b$$
$$E(X + Y) = E(X) + E(Y)$$
$$\text{En general}: E\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} E(X_i)$$

De hecho, vemos que para el caso discreto

$$\begin{aligned} E(aX + b) &= \sum_{x} (aX + b)\, p(x) \\ &= a\sum_{x} xp(x) + b\sum_{x} p(x) \\ &= aE(X) + b \end{aligned}$$

ii) $E(a)=a$

iii) $E(X \cdot Y)=E(X) \cdot E(Y)$ si *X, Y* son independientes

iv) Momentos de orden *K* de una v.a.

$$E(X^k) = \begin{cases} \sum_{W} x_i^k p(x_i) & \text{(v.a. discreta)} \\ \int_{W} x^k f(x)dx & \text{(v.a. continua)} \end{cases}$$

v) $E(X)$ es una constante, i.e. :

$E(E(X))=E(X)$

III) Varianza

La varianza es una medida de la dispersión de los datos, es, decir, que tan lejos están unos de otros. A la raíz cuadrada de la varianza se le denomina desviación estándar.

Definición:

$$Var(X) = E\left[(X - E(X))^2\right] = E(X^2) - [E(x)]^2$$

Dem.:

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

a) Propiedades

Sean X, Y v.a , a y b constantes

i) $Var(aX+b) = a^2 Var(X)$

Dem.:

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

ii) $Var(a) = 0$

iii) Sean $X_1,\ldots,X_n$ v.a independientes, entonces: $Var\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} Var(X_i)$

En general, si $X_1,\ldots,X_n$ NO son independientes:

$$Var\left(\sum_{i=1}^{n}\alpha_i X_i\right)=\sum_{i=1}^{n}\alpha_i^2 Var(X_i)+\sum_i\sum_j \alpha_i\alpha_j Cov(X_i,X_j)$$

Donde: $Cov(X,Y) = E[(X-E(X))\cdot(Y-E(Y))] = E(X\cdot Y) - E(X)\cdot E(Y)$

NOTA: $Cov(X,X) = Var(X)$

Si X, Y son v.a independientes, $E(X\cdot Y) = E(X)\cdot E(Y)$

$\Rightarrow Cov(X,Y) = 0 \Rightarrow Var(X+Y)=Var(X)+Var(Y)$

iv) $Var(X)\geq 0$

v) Desviación Estándar = $\sqrt{Var(X)}$

IV) Función Generatriz de Momentos (f.g.m.)

$$M_x(t)=E(e^{tX})=\begin{cases}\sum_{W} e^{tx_i}p(x_i) & \text{(v.a. discreta)}\\ \int_{W} e^{tx}f(x)dx & \text{(v.a. continua)}\end{cases}$$

a) Propiedades

i) $M_x^{(n)}(0)=E(X^n)$

debido a que,

$$M'(t)=\frac{d}{dt}E\left(e^{tX}\right)=E\left[\frac{d}{dt}\left(e^{tX}\right)\right]=E\left[Xe^{tX}\right]$$

Así, $M_X'(0)=E(X)$

De manera similar:

$$M_X''(t)=\frac{d}{dt}M_X'(t)=\frac{d}{dt}E\left[Xe^{tX}\right]=E\left[\frac{d}{dt}\left(Xe^{tX}\right)\right]=E\left[X^2e^{tX}\right]$$

$$\Rightarrow M_X''(0)=E\left(X^2\right)$$

En general: $M_X^{(n)}(0)=E\left[X^n\right]$

ii) Si $Y = aX + b$ => $M_Y(t) = e^{bt}M_x(at)$

iii) Si $Z = X + Y$, X e Y son independientes: $M_Z(t) = M_X(t) \cdot M_Y(t)$

Dem.:

$$
\begin{aligned}
M_z(t) &= M_{X+Y}(t) \\
&= E\left[e^{t(X+Y)}\right] \\
&= E\left[e^{tX}e^{tY}\right] \\
&= E\left[e^{tX}\right]E\left[e^{tY}\right] \\
&= M_X(t)M_Y(t)
\end{aligned}
$$

En general, si $X_1,\ldots,X_n$ son v.a independientes, tales que:

$$Z = \sum_{i=1}^{n} X_i$$

$$M_z(t) = \prod_{i=1}^{n} M_{X_i}(t)$$

iv) Teorema: Si $M_X(t) = M_Y(t)$ => $X = Y$

v) f.d.p. de una función de una v.a. continua X:

Sea $Y = H(X)$ una función de X, luego: $g(y) = f(H^{-1}(y)) \cdot \left|\dfrac{dH^{-1}(y)}{dy}\right|$

v) Independencia

Sean dos variables aleatorias X e Y, estadísticamente independientes, entonces la densidad conjunta es igual al producto de sus densidades marginales:

$$f(x, y) = f_X(x) * f_Y(y)$$

O de otra forma, $F(x, y) = F_X(x) * F_Y(y) = P(X \le x) * P(Y \le y)$

Además,

- $E(X * Y) = E(X) * E(Y)$
- $Cov(X, Y) = 0$
- Si $Z = X + Y$, entonces $M_Z(t) = M_X(t) * M_Y(t)$

- En general, si $X_1, X_2 ..., X_n$ son v.a.s independientes tal que $Z = \sum_{i=1}^{n} X_i$, entonces $M_Z(t) = \prod_{i=1}^{n} M_{X_i}(t)$

vi) Ley de los Grandes Números

Suponga que $X_1, X_2 ..., X_n$ es una secuencia de variables aleatorias i.i.d. con media μ y varianza σ^2 , y sea $S_n = X_1 + X_2 + ... + X_n$, de modo que $E(S_n) = n\mu$ y $Var(S_n) = n\sigma^2$, entonces

- Ley débil de los grandes números muestra que para $\varepsilon > 0$

$$P(|S_n / n - \mu| > \varepsilon) \to 0 \quad \text{cuando} \quad n \to \infty$$

Dem.:

$$E\left[\frac{S_n}{n}\right] = \mu \qquad \text{y} \qquad Var\left[\frac{S_n}{n}\right] = \frac{\sigma^2}{n}$$

Luego por la desigualdad de Chebyshev

$$P\left[\left|\frac{S_n}{n} - \mu\right| > \varepsilon\right] \le \frac{\sigma^2}{n\varepsilon^2} \xrightarrow[n\to\infty]{} 0$$

Nota:
Chebyshev:
Si X es una v.a. con media μ y varianza σ^2, entonces $\forall k > 0$

$$P(|X - \mu| \ge k) \le \frac{\sigma^2}{k^2}$$

- Ley fuerte de los grandes números muestra que

$$P(S_n / n \to \mu) \to 1$$

vii) Teorema Central del Límite

Suponga que $X_1, X_2 ..., X_n$ es una secuencia de variables aleatorias i.i.d. con media μ y varianza σ^2, y sea $S_n = X_1 + X_2 + ... + X_n$, de modo que $E(S_n) = n\mu$ y $Var(S_n) = n\sigma^2$, entonces para un número n "suficientemente grande".

$$\Rightarrow \frac{S_n - n\mu}{\sigma\sqrt{n}} \sim N(0,1)$$

Este resultado es muy importante en estadística, ya que para un tamaño de muestra grande, la suma o promedio de variables aleatorias i.i.d. sigue una distribución Normal, sin importar la distribución de origen de las variables de la suma.

En términos generales, $$\frac{X - E(X)}{\sqrt{Var(X)}} \sim N(0,1)$$

Problemas

Problema 1:

Sea $\beta = e^{-\lambda \overline{X}_n}$ y $X \sim f(X) = \frac{1}{\sigma}\exp\left(\frac{-x}{\sigma}\right)$. Calcule la esperanza y la varianza de β (Se recomienda usar la función generadora de momento de X)

Respuesta

$$\mathrm{E}(\beta) = \mathrm{E}\left(e^{-\lambda \overline{X}_n}\right) = \mathrm{E}\left(e^{-\frac{\lambda}{n}\sum X_i}\right)$$

Por la indicación recordamos que $M_x(t) = \mathrm{E}\left(e^{tx}\right)$ reconociendo términos $t = -\frac{\lambda}{n}$ y $X = \sum X_i$

también sabemos que $M_X(t) = M_{\sum X_i}(t) = \prod M_{X_i}(t)$

$M_{X_i}(t) = \frac{1}{1-\sigma t}$ (desarrollado en la clase auxiliar)

$$\Rightarrow \mathrm{E}\left(e^{-\frac{\lambda}{n}\sum X_i}\right) = M_X(t) = M_{\sum X_i}\left(-\frac{\lambda}{n}\right) = \left(\frac{1}{1+\sigma\frac{\lambda}{n}}\right)^n = \left(\frac{n}{n+\sigma\lambda}\right)^n$$

$$\mathrm{var}(\beta) = \mathrm{E}\left(\beta^2\right) - \mathrm{E}(\beta)^2$$

$$\mathrm{E}\left(\beta^2\right) = \mathrm{E}\left(e^{-2\frac{\lambda}{n}\sum X_i}\right) = \left(\frac{1}{1+\frac{2\lambda\sigma}{n}}\right)^n = \left(\frac{n}{n+2\sigma\lambda}\right)^n$$

$$\mathrm{var}(\beta) = \left(\frac{n}{n+2\sigma\lambda}\right)^n - \left(\frac{n}{n+\sigma\lambda}\right)^{2n}$$

Problema 2: La producción mínima de una máquina es de 2 mil tornillos diarios y la máxima es de 6 mil. Si la densidad de probabilidad del número x de miles de tornillos se puede representar por la función:

$$f(x) = \frac{3}{128}(8x^2 - x^3 - 12x)$$

Calcule
a.- La producción más probable
b.- La varianza de la cantidad producida
c.- La función distribución acumulada (FDA)

Solución:

a.- La producción más probable viene dada por el punto donde se maximiza la función de probabilidad.

$$f'(x) = \frac{3}{128}(16x - 3x^2 - 12) = 0 \Rightarrow 16x - 3x^2 - 12 = 0$$

Reconocemos una ecuación de segundo grado dónde $x_1 = 4.47$ y $x_2 = 0.87$. Cómo la producción fluctúa entre 2mil y 6 mil tornillos diarios x_1 es la producción más probable.

b.- $Var(x) = E(x^2) - E(x)^2$

$$E(x) = \frac{3}{128}\int_2^6 x(8x^2 - x^3 - 12x)dx = \frac{3}{128}\int_2^6 (8x^3 - x^4 - 12x^2)dx = 4.2$$

$$E(x^2) = \frac{3}{128}\int_2^6 x^2(8x^2 - x^3 - 12x)dx = \frac{3}{128}\int_2^6 (8x^4 - x^5 - 12x^3)dx = 18.4$$

Con lo que

$$Var(x) = 18.4 - (4.2)^2 = 0.76$$

c.- FDA en todo el espacio.

c.1.- x > 6.000

$$F(x) = P(X \le x) = 1$$

c.2.- x < 2.000

$$F(x) = P(X \le x) = 0$$

c.3 2.000<X<6.000

$$F(x) = P(X \le x) = \int_2^x f(x)dx = \frac{3}{128}\left[\frac{8}{3}x^3 - \frac{1}{4}x^4 - 6x^2 + \frac{20}{3}\right]$$

Cualquier error escribir a gonmat82@yahoo.com