EYP1027 Modelos Probabilísticos

Clase 8

Profesor: Reinaldo B. Arellano-Valle

Departamento de Estadística
Pontificia Universidad Católica de Chile

Primer Semestre 2020

Contenido I

Formas alternativas de caracterizar una distribución o variable aleatoria son las denominadas funciones generadoras, las cuales también permiten generar los momentos de la distribución cuando existen.

Definición 1.1

Función generadora de momentos La función generadora de momento (fgm) de una variable aleatoria X, se define como,

$$\begin{aligned} M_X(t) &= \mathrm{E}\left(e^{tX}\right) \\ &= \begin{cases} \sum_x e^{tx} f_X(x) & \text{si } X \text{ es discreta,} \\ \int_{-\infty}^{\infty} e^{tx} f_X(x) dx & \text{si } X \text{ es continua,} \end{cases} \end{aligned}$$

siempre que la respectiva esperanza exista (sea finita) para $|t| < h$, para algún $h > 0$. Si dicha esperanza no existe en una vecindad de 0, se dice que la fgm no existe.

Nota: La función $C_X(t) = \log M_X(t)$ se llama función generadora de cumulantes.

Teorema 1.1

Sea X una variable aleatoria cuya fgm, $M_X(t)$, existe en una vecindad de 0. Entonces, se tiene que:

i) Para todo $k = 1, 2, \ldots$, $\mathrm{E}(X^k)$ es finito y puede calcularse como

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

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

ii) Sea Y otra variable aleatoria cuya fgm existe. Si $M_Y(t) = M_X(t)$ para todo t en una vecindad de 0, entonces $F_Y(z) = F_X(z)$ para todo z.

iii) Si $Y = aX + b$, entonces $M_Y(t) = e^{tb} M_X(at)$.

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 $\mathrm{E}(X^k)$ existe para todo $k = 1, 2, \ldots$. Luego, $M_X(t)$ existe para todo t, y esta dada por

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

Sin embargo, al usar que $e^t = \sum_{k=0}^{\infty} \frac{t^k}{k!}$ para todo t, se tiene que

$$\begin{aligned} M_X(t) &= \frac{1}{t^2}\left(\sum_{k=0}^{\infty} \frac{t^k}{k!} + \sum_{k=0}^{\infty} \frac{(-t)^k}{k!} - 2\right) \\ &= \frac{1}{t^2}\left(\sum_{k=0}^{\infty} \frac{t^k + (-t)^k}{k!} - 2\right) \\ &= \frac{1}{t^2}\left(\sum_{k=1}^{\infty} \frac{2t^{2k}}{(2k)!}\right) \\ &= 2\sum_{k=1}^{\infty} \frac{t^{2k-1}}{(2k)!}, \end{aligned}$$

de donde se concluye que $\mathsf{E}(X^{2k-1}) = 0$ y $\mathsf{E}(X^{2k}) = \frac{2}{(2k+1)(2k+2)}$, $k = 1, 2, \ldots$.

Si X es una variable aleatoria en $(\Omega, \mathcal{A}, P)$, también se puede considerar la función generadora definida por,

$$G_X(t) = \mathsf{E}(t^X),$$

siempre que la esperanza exista. Puede notarse facílmente que,

a) Si $M_X(t)$ existe en una vecindad de 0, entonces, $G_X(t) = M_X(\log t)$, $t > 0$.

b) Si X es una variable aleatoria discreta con recorrido $\mathcal{X} = \{0, 1, 2, \ldots\}$, entonces $G_X(t) = \sum_{x=0}^{\infty} t^x P(X = x)$ y se llama función generadora de probabilidad (fgp). En tal caso, $G_X(t)$ converge al menos para $|t| \leq 1$, ya que $G_X(1) = \sum_{x=0}^{\infty} P(X = x) = 1$.

c) Si $\mathsf{E}(|X|^k) < \infty$, entonces, $E\{X(X-1)\cdots(X-k+1)\} = G_X^{(k)}(1)$, donde $G_X^{(k)}(1) = \frac{d^k}{dt^k} G_X(t)\,|_{t=1}$

La fgp juega un papel importante en muchas áreas de probabilidad aplicada, en particular, en procesos estocásticos. Una definición más formal de la fgp es dada a continuación.

Definición 1.2

Función generatriz de probabilidad Sea X una variable aleatoria no negativa con valores enteros y sea $p_k = P(X = k)$, $k = 0, 1, 2, \cdots$, con $\sum_{k=0}^{\infty} p_k = 1$. La fgp de X se define como,

$$G_X(t) = E(t^X) = \sum_{k=0}^{\infty} p_k t^k, \quad |t| < 1.$$

Ejemplo 1.2

Sea X una variable aleatoria con fmp dada por

$$f_X(x) = P(X = x) = \begin{cases} \binom{n}{x} p^x (1-p)^{n-x}, & \text{si } x = 0, 1, \ldots, n, \\ 0, & \text{eoc}, \end{cases}$$

donde $n \in \{1, 2, \ldots\}$ y $p \in (0, 1)$. Entonces,

$$\begin{aligned} G_X(t) &= \sum_{x=0}^{n} t^x \binom{n}{x} p^x (1-p)^{n-x} \\ &= (1-p)^n \sum_{x=0}^{n} \binom{n}{x} \left(\frac{pt}{1-p} \right)^x \\ &= (1-p)^n \left(\frac{pt}{1-p} + 1 \right)^n = (pt + 1 - p)^n, \quad t > 0. \end{aligned}$$

Debido a que las funciones generadoras previas no siempre existen, se define la función carasterística con recorrido en los números complejos.

Definición 1.3

Función característica La función característica (fc) de una variable aleatoria X se define definida como,

$$\varphi_X(t) = \mathrm{E}(e^{itX}) = \mathrm{E}(\cos tX) + i\mathrm{E}(\sin tX),$$

donde $i = \sqrt{-1}$, es decir, $\varphi_X : \mathbb{R} \longrightarrow \mathbb{C}$.

Nota: Como $|E(e^{itX})| \leq E(|e^{itX}|) = \mathsf{E}(\sqrt{\cos^2 tX + \sin^2 tx}) = \mathsf{E}(1) = 1$, entonces $\varphi_X(t)$ existe para todo t y para toda variable aleatoria X.

Propiedades básicas:

a) $|\varphi_X(t)| \leq \varphi_X(0) = 1$ para todo t.
b) $\overline{\varphi_X(t)} = \mathsf{E}(e^{-itX}) = \varphi_X(-t)$.
c) $\varphi_X(t)$ es uniformente continua.
d) $\varphi_X(t) = \varphi_Y(t)$ para todo $t \iff X \stackrel{d}{=} Y \iff F_X(z) = F_Y(z)$ para todo z.
e) $\varphi_X(t) =$ función real $\iff X \stackrel{d}{=} -X$ (simetría con respecto de 0).
f) $\varphi_{aX+b}(t) = e^{ibt} = \varphi_X(at)$.
g) Si $\mathsf{E}(|X|^k) < \infty \implies \varphi_X^{(k)}(0) = i^k \mathsf{E}(X^k)$, $k = 1, 2, \ldots$
h) Si $M_X(t)$ existe en una vecindad de 0, entonces $\varphi_X(t) = M_X(it)$ o $M_X(t) = \varphi_X(-it)$.

Ejemplo 1.3

1. Sea X una variable aleatoria con, $P_X(X=-1) = P_X(X=1) = 1/2$. Entonces,

$$\varphi_X(t) = \frac{1}{2}(\cos t + \cos(-t)) + \frac{i}{2}(\sin t + \sin(-t))$$
$$= \cos t \implies X \stackrel{d}{=} -X.$$

2. Sea X una variable aleatoria con fdp $f_X(x) = \frac{1}{\pi(1+x^2)}$ $x \in \mathbb{R}$. Entonces,

$$\varphi_X(t) = \int_{-\infty}^{\infty} \frac{e^{itx}}{\pi(1+x^2)} dx = e^{-|t|}.$$

En este caso, la fgm de X no existe.

Teorema 1.2

Regla de Leibnitz Si $f(x,\theta), a(\theta)$, y $b(\theta)$ son diferenciables con respecto a θ, entonces

$$\frac{d}{d\theta}\int_{a(\theta)}^{b(\theta)} f(x,\theta)dx = f(b(\theta),\theta)\frac{d}{d\theta}b(\theta) - f(a(\theta),\theta)\frac{d}{d\theta}a(\theta) + \int_{a(\theta)}^{b(\theta)} \frac{\partial}{\partial\theta} f(x,\theta)dx.$$

Note que si $a(\theta)$ y $b(\theta)$ son constantes, tenemos el caso especial,

$$\frac{d}{d\theta}\int_{a}^{b} f(x,\theta)dx = \int_{a}^{b} \frac{\partial}{\partial\theta} f(x,\theta)dx.$$

Distribuciones comunes

Este capítulo cataloga varias de las distribuciones estadísticas más comunes. Para cada distribución se entrega su media, varianza y otras medidas descriptivas útiles. También se indican algunas aplicaciones típicas y algunas interrelaciones útiles. Este capítulo no es de ninguna manera completo en su cobertura de distribuciones estadísticas, particularmente considerando que en las últimas decadas han surgido una enormidad de nuevas distribuciones de probabilidad.

Nota: Para simplificar la notación, y cuando no sea necesario, se eliminaran todos los subíndices; $F_X = F$, $f_X = f$, $M_X = M$, $\mu_X = \mu$, $\sigma_X^2 = \sigma^2$, etc.

Distribuciones discretas

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,

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

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

Teorema 2.1

Si $X \sim UD(\{x_1, ..., x_N\})$, entonces,

i) $\mathrm{E}(X^r) = \frac{1}{N} \sum_{i=1}^{N} x_i^r, \quad k = 1, 2, \ldots$

ii) $M_X(t) = \frac{1}{N} \sum_{i=1}^{N} e^{tx_i}$.

Distribuciones discretas

Corolario 2.1

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

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

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

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

Demostración 2.1

Basta recordar que,

$$\sum_{x=1}^{N} x = \frac{N(N+1)}{2} \quad \text{y} \quad \sum_{x=1}^{N} x^2 = \frac{N(N+1)(2N+1)}{6}.$$

Ejemplo 2.1

Sea X el número obtenido al lanzar un dado equilibrado. El espacio muestral es $\mathcal{X} = \{1, 2, 3, 4, 5, 6\}$ y cada resultado tiene probabilidad $p(x) = \frac{1}{6}$. Por lo tanto, $X \sim UD(\{1, 2, 3, 4, 5, 6\})$, distribución uniforme discreta tal que,

$$
\begin{aligned}
p(x) &= \frac{1}{6}, \quad \text{si} \quad x = 1, 2, 3, 4, 5, 6, \\
\mathrm{E}(X) &= \frac{1+2+3+4+5+6}{6} = \frac{6+1}{2} = 3.5, \\
\mathrm{Var}(X) &= \frac{6^2 - 1^2}{12} = \frac{35}{12}.
\end{aligned}
$$

Tarea 2.1

Calcule la fgm de X, $M(t) = \mathrm{E}(e^{tX})$.

Distribuciones discretas

Distribución Binomial La distribución binomial, una de las distribuciones discretas más útiles, se basa en la idea de un ensayo de Bernoulli. El ensayo de Bernoulli es un experimento con dos, y sólo dos, resultados posibles, $\Omega = \{E, F\}$, con $E =$ éxito, $F =$ fracazo, $P(E) = p$, $0 < p < 1$.

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.

Notación: $X \sim Bin(n, p)$ y $X \sim Ber(p)$ si $n = 1$.

Teorema 2.2

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

i) $\mathrm{E}(X) = np$

ii) $\mathrm{Var}(X) = npq$,

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

Demostración 2.2

En este caso tenemos,

$$\begin{aligned} M(t) &= \sum_{x=0}^{n} e^{tx} \binom{n}{x} p^x (1-p)^{n-x} \\ &= \sum_{x=0}^{n} \binom{n}{x} \left(pe^t\right)^x q^{n-x} \\ &= \left(pe^t + q\right)^n . \end{aligned}$$

Ejemplo 2.2

Una prueba de selección multiple consta de 10 preguntas y cada pregunta tiene 4 respuesta de las cuales sólo una es la correcta. El rendimiento de un postulante se considera satisfactorio si obtiene por lo menos 6 repuestas correctas. Un postulante no preparado elige las respuestas al azar. Calcule la probabilidad de que su rendimiento sea considerado satisfactorio.
Sean $A = \{\text{Se elige la respuesta correcta}\}$ y $p = P(A)$, y defina $X =$ número de respuestas correctas (#de veces que ocurre A). Aquí, $n = 10$, $p = 0.25$, $\mathcal{X} = \{0, 1, ..., 10\}$, y se tiene que, $X \sim Bin(10, 0.25)$.

Se pide $\mathrm{P(X{\geq}6)} = \sum_{x=6}^{10} \binom{10}{x} (0.25)^x (0.75)^{10-x} = 0.20$

Distribución Hipergeométrica La distribución hipergeométrica tiene muchas aplicaciones en el muestreo de poblaciones finitas.

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,

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

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

Notación: $X \sim Hip(n, K, N)$.

Ejemplo 2.3

Una compañía recibe una partida de 20 items. Debido a que la inspección de cada items es costosa, se chequea al azar 6 artículos. La partida se acepta si no más uno de los items inspeccionados es defectuoso. Calcule la probabilidad de que una partida con 5 item defectuosos sea aceptada.
Sea $A =$ el artículo elegido es defectuoso, y defina la variable aleatoria $X =$ número de artículos defectuosos obtenidos en la muestra
Luego, $X \sim Hip(6, 5, 20)$, y

$$p(x) = \begin{cases} \frac{\binom{5}{x}\binom{15}{6-x}}{\binom{20}{6}}, & x = 0, 1, 2, 3, 4, 5, \\ 0, & \text{eoc.} \end{cases}$$

Se pide $\quad \mathrm{P}(X \leq 1) \ = \ p(0) + p(1) \ = \ 0.129 + 0.387 \ = \ 0.516.$

Ejemplo 2.4

Se embarcan motores eléctricos pequeños en lotes de 50. Antes de que tal cargamento sea aceptado, un inspector elige 5 motores y los inspecciona. Si ninguno de los motores probados es defectuoso, el lote se acepta. Si se encuentra que uno o más son defectuosos, se inspecciona el cargamento completo. Supongamos que, en realidad, hay tres motores defectuosos en el lote. Cuál es la probabilidad de que sea necesaria una inspección de 100 % ?
Si denotamos por X el número de motores defectuosos encontrados, se necesitará una inspección de 100 % si y sólo si $X \geq 1$. Luego,

$$P(X \geq 1) = 1 - P(X = 0) = 1 - \frac{\begin{pmatrix} 3 \\ 0 \end{pmatrix}\begin{pmatrix} 47 \\ 5 \end{pmatrix}}{\begin{pmatrix} 50 \\ 5 \end{pmatrix}} = 0.28$$

Distribuciones discretas

Teorema 2.3

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

i) $\mathrm{E}(X) = np$

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

iii) $p(x) \simeq \left(\begin{array}{c} n \\ x \end{array} \right) p^x (1 - p)^{n - x}$, para N grande.

Demostración 2.3

Se deja al lector. Ver Casella & Berger (2002).

Ejemplo 2.5

En una ciudad con 2 millones de habitantes, 60 % pertenecen al partido político A. Cien personas son elegidas al azar de los habitantes. La distribución del número de personas, entre las 100 elegidas que pertenecen al partido A, es una variable aleatoria X con distribución $Hip(100, 120000, 2000000)$. Aplicando el resultado anterior podemos aproximar esta distribución por una binomial con parámetros, $n = 100$ y $p = 0.6$. De esa manera, por ejemplo, para la probabilidad de que entre las 100 personas elegidas exactamente 40 pertenezcan al partido A, se tiene que

$$P(X = 40) \simeq \binom{100}{40} (0.6)^{40}(0.4)^{60} = 2.4425 \times 10^{-5}$$

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,

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

Notación: $X \sim P(\lambda)$

Teorema 2.4

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

i) $\mathrm{E}(X) = \lambda$

ii) $\mathrm{Var}(X) = \lambda$

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

Demostración 2.4

Calculemos la fgm. En efecto, tenemos que,

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

Por lo tanto,

$$
\begin{aligned}
E(X) &= \left.\frac{d}{dt}M(t)\right|_{t=0} = \left.\lambda e^t \exp\left\{\lambda\left(e^t-1\right)\right\}\right|_{t=0} \\
&= \lambda, \text{ y} \\
E\left(X^2\right) &= \left.\frac{d^2}{dt^2}M(t)\right|_{t=0} \\
&= \left.\left[\lambda e^t \exp\left\{\lambda\left(e^t-1\right)\right\} + \lambda^2 e^{2t}\exp\left\{\lambda\left(e^t-1\right)\right\}\right]\right|_{t=0} \\
&= \lambda(\lambda+1),
\end{aligned}
$$

de donde sigue que $\operatorname{Var}(X) = \lambda(\lambda+1) - \lambda^2 = \lambda$.

Ejemplo 2.6

El número de pacientes que acuden diariamente a la sala de emergencia (ER) de cierto hospital tiene una distribución de Poisson con una media de 10 personas/dia. Cuál es la probabilidad de que, durante un día normal, el número de pacientes admitidos en la sala de emergencia del hospital sea menor o igual a 3 ?
Sea $X = \{\text{cantidad de pacientes que acuden diariamente a la ER}\}$. Del enunciado del problema, se tiene que $X \sim P(10)$, y por lo tanto,

$$\begin{aligned} P(X \leq 3) &= P(X = 0) + P(X = 1) + P(X = 2) + P(X = 3) \\ &= e^{-10} + 10e^{-10} + 50e^{-10} + \frac{1000}{6}e^{-10} \\ &= 1.0336 \times 10^{-2}. \end{aligned}$$

Teorema 2.5

Sea X una variable aleatoria con distribución binomial con parámetros n y p. Esto es,

$$P(X = x) = \binom{n}{x} p^x (1-p)^{n-x}, \quad x = 0, 1, \ldots, n.$$

Supóngase que cuando $n \to \infty, np = \lambda$ (constante), o equivalente, cuando $n \to \infty, p \to 0$ tal que $np \to \lambda$. Entonces, bajo estas condiciones se tiene que,

$$\lim_{n \to \infty} P(X = x) = \frac{e^{-\lambda} \lambda^x}{x!}, \quad x = 0, 1, \ldots,$$

la distribución de Poisson con parámetro λ.

Este teorema, esencialmente, dice que se puede aproximar las probabilidades binomiales con las probabilidades de la distribución de Poisson siempre que n sea grande y p pequeño.

Demostración 2.5

Sea $X \sim Bin(n,p)$, y sea $np = \lambda$, de donde $p = \lambda/n$. Entonces,

$$
\begin{aligned}
P(X=x) &= \left(\begin{array}{c} n \\ x \end{array}\right) p^x (1-p)^{n-x} \\
&= \frac{n!}{x!(n-x)!}\left(\frac{\lambda}{n}\right)^x \left(1-\frac{\lambda}{n}\right)^{n-x} \\
&= \frac{\lambda^x}{x!}\,\frac{n!}{n^x(n-x)!}\left(1-\frac{\lambda}{n}\right)^{n-x} \\
&= \frac{\lambda^x}{x!}\left(1-\frac{\lambda}{n}\right)^n \frac{n(n-1)\cdots(n-x+1)}{n^x}\left(1-\frac{\lambda}{n}\right)^{-x} \\
&= \frac{\lambda^x}{x!}\left(1-\frac{\lambda}{n}\right)^n \left[(1)\left(1-\frac{1}{n}\right)\cdots\left(1-\frac{x-1}{n}\right)\right]\left(1-\frac{\lambda}{n}\right)^{-x}.
\end{aligned}
$$

Aquí, se tiene que,

i) Para cada $x \geq 0$,

$$\lim_{n\to\infty} \frac{n}{n} \cdot \frac{n-1}{n} \cdot \dots \cdot \frac{n-x+1}{n} \left(1 - \frac{\lambda}{n}\right)^{-x} = 1,$$

ii) Es conocido (de la definición del número e) que,

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

$\therefore \lim_{n\to\infty} P(X = x) = e^{-\lambda}\lambda^x/x!$; es decir, en el límite se obtiene la distribución de Poisson con parámetro λ.

La siguiente tabla muestra qué tan buena es la aproximación cuando $\lambda = np = 1$.

x	$P(\lambda)$	$Bin(100, 0.01)$	$Bin(10, 0.1)$
0	0.3678	0.3660	0.3487
1	0.3679	0.3697	0.3874
2	0.1839	0.1848	0.1937
3	0.0613	0.0610	0.0574
4	0.0153	0.0149	0.0112

Empíricamente, se ha determinado que la aproximación puede usarse de manera segura si $n \geq 100, p \leq 0.01$ y $np \leq 20$

Ejemplo 2.7

Hay 135 estudiantes dentro de una sala de conferencias. La probabilidad de que uno de los estudiantes celebre su cumpleaños hoy es igual a $\frac{1}{365}$. Cuál es la probabilidad de que dos o más estudiantes en la sala de conferencias estén celebrando sus cumpleaños hoy ? Sea $X =$ número de estudiantes que celebran sus cumpleaños hoy. Se sabe que $X \sim Bin(135, \frac{1}{365})$. Sin embargo esta distribución puede ser apróximada por una distribución de Poisson con parámetro $\lambda = 135/365 = 27/73$. Por lo tanto,

$$
\begin{aligned}
P(X \geq 2) &= 1 - P(X = 0) - P(X = 1) \\
&= 1 - e^{-\frac{27}{73}} - \frac{27}{73} e^{-\frac{27}{73}} \\
&= 5.3659 \times 10^{-2}.
\end{aligned}
$$

Distribuciones discretas

Distribuciones binomial negativa y geométrica

La distribución binomial cuenta el número de éxitos en un número fijo de ensayos de Bernoulli. Si, en cambio, se cuenta el número de ensayos de Bernoulli necesarios para obtener exactamente r éxitos, entonces se tienen las distribuciones binomial negativa $(r \geq 1)$ y geométrica $(r = 1)$. En la derivación de estas distribuciones se usa la distribución binomial. En efecto, el evento $\{X = x\}$ puede ocurrir sólo si hay exactamente $r - 1$ éxitos en los primeros $x - 1$ ensayos, y un éxito en el ensayo $x-$ésimo, es decir, el resultado del experimento debe tener la forma:

$$(\underbrace{FF\cdots F \underbrace{E}_{\text{1er éxito}} FF\cdots F \underbrace{E}_{\text{2do éxito}} FF\cdots F \quad \underbrace{E}_{(r-1)\text{-ésimo éxito}} \quad FF\cdots F}_{x-1 \text{ ensayos}} \underbrace{E}_{r\text{-ésimo éxito}})$$

Entonces, la probabilidad de $r-1$ éxitos en $x-1$ ensayos es la probabilidad binomial

$$\left(\begin{array}{c} x-1 \\ r-1 \end{array} \right) p^{r-1}(1-p)^{x-r},$$

y con probabilidad p hay un éxito en el $x-$ésimo ensayo. Multiplicando estas probabilidades se obtiene

$$P(X=x) = \left(\begin{array}{c} x-1 \\ r-1 \end{array} \right) p^{r}(1-p)^{x-r},$$

para $x = r, r+1, \ldots$, ya que al menos se requieren r ensayos para obtener r éxitos.

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,

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

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

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

Notación: Binomial negativa, $X \sim BN(r, p)$. Geométrica, $X \sim Geo(p)$.

Ejemplo 2.8

En un departamento de control de calidad, se inspeccionan las unidades que provienen de una línea de ensamblaje. Si la proporción de unidades defectuosas es 0.03, cuál es la probabilidad de que la vigésima unidad inspeccionada sea la tercera defectuosa ?
Sea X = número de unidades necesarias a inspeccionar para obtener exactamente tres defectuosas. Luego tenemos que $X \sim BN(3, 0.03)$.
Por lo tanto,

$$\begin{aligned} P(X = 20) &= \binom{19}{2} (0.03)^3 (1 - 0.03)^{17} \\ &= 2.7509 \times 10^{-3}. \end{aligned}$$

Teorema 2.6

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

i) $\mathrm{E}(X) = r/p$

ii) $\mathrm{Var}(X) = r(1-p)/p^2$

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

Demostración 2.6

El $k-$ésimo momento de X alrededor del origen está dado por,

$$
\begin{aligned}
\mathrm{E}(X^k) &= \sum_{j=r}^{\infty} j^k \left(\begin{array}{c} j-1 \\ r-1 \end{array} \right) p^r (1-p)^{j-r} \\
&= \frac{r}{p} \sum_{j=r}^{\infty} j^{k-1} \left(\begin{array}{c} j \\ r \end{array} \right) p^{r+1} (1-p)^{j-r} = \frac{r}{p} \mathrm{E}\left\{ (Y-1)^{k-1} \right\},
\end{aligned}
$$

donde $Y \sim BN(r+1,p)$. Por lo tanto,

$$\mathrm{E}(X) = \frac{r}{p} \quad \text{y} \quad \mathrm{E}(X^2) = \frac{r}{p}\,\mathrm{E}(Y-1) = \frac{r}{p} \times \frac{r+1-p}{p},$$

lo cual produce,

$$\begin{aligned} \mathrm{Var}(X) &= \frac{r^2+r-rp}{p^2} - \frac{r^2}{p^2} \\ &= \frac{r(1-p)}{p^2}. \end{aligned}$$

Por otro lado,

$$M(t) = \sum_{j=r}^{\infty} e^{tj} \begin{pmatrix} j-1 \\ r-1 \end{pmatrix} p^r (1-p)^{j-r} = \left(pe^t\right)^r \left\{1-(1-p)e^t\right\}^{-r};$$

ver Casella & Berger (2002), para más detalles.

Corolario 2.2

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

i) $\mathrm{E}(X) = 1/p$

ii) $\mathrm{Var}(X) = (1-p)/p^2$

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

Para la variable aleatoria $Y = X - r$, se tiene que,

$$E(Y) = \frac{r}{p} - r = \frac{r(1-p)}{p}$$

$$\mathsf{Var}(Y) = \mathsf{Var}(X) = \frac{r(1-p)}{p^2}$$

$$M_Y(t) = \left\{\frac{p}{1-(1-p)e^t}\right\}^r.$$

Distribuciones discretas

La distribución geométrica tiene una propiedad interesante, conocida como la propiedad "sin memoria". Para enteros $s > t$, se tiene que,

$$P(X > s|X > t) = P(X > s - t), \tag{2.1}$$

es decir, la distribución geométrica "olvida" lo ocurrido en el pasado. La probabilidad de obtener $s - t$ fallas adicionales, después de haber observado t fallas, es la misma que la probabilidad de observar $s - t$ fallas al comienzo de la secuencia. En otras palabras, la probabilidad de obtener una serie de fallas depende solo de la duración de la ejecución, no de su posición. En efecto,

$$\begin{aligned} P(X > s|X > t) &= \frac{P(X > s \text{ y } X > t)}{P(X > t)} \\ &= \frac{P(X > s)}{P(X > t)} \\ &= (1-p)^{s-t} = P(X > s - t). \end{aligned}$$

Ejemplo 2.9

Tiempos de falla La distribución geométrica a veces se usa para modelar tiempo hasta la falla de ciertos componentes. Por ejemplo, si la probabilidad es 0.001 de que una bombilla falle en un día determinado, entonces la probabilidad de que dure al menos 30 días es

$$P(X > 30) = \sum_{x=31}^{\infty} 0.001(1-0.001)^{x-1} = (0.999)^{30} = 0.970.$$

La propiedad sin memoria de la distribución geométrica describe una propiedad muy especial de "falta de envejecimiento". Indica que la distribución geométrica no es aplicable al modelado de tiempos de falla, para las cuales se espera que la probabilidad de falla aumente con el tiempo.

Distribuciones continuas

En esta sección discutiremos algunas de las familias más comunes de distribuciones continuas, aquellas con nombres conocidos. Las distribuciones mencionadas aquí de ninguna manera constituyen todas las distribuciones utilizadas en estadística. De hecho, como ya fue mencionado, cualquier función integrable no negativa se puede transformar en una pdf.

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,

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

Notación: $X \sim U(a, b)$ (o $X \sim U[a, b]$).

Teorema 2.7

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

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

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

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

Demostración 2.7

Se deja como ejercicio para el lector.

Ejemplo 2.10

Sea $X \sim U(-3, 2)$. Calcule $P(X \geq 0)$ y $P(-5 \leq X \leq \frac{1}{2})$. En este caso la fdp de X está dada por,

$$f(x) = \begin{cases} \frac{1}{5}, & \text{si } -3 < x < 2 \\ 0, & \text{eoc} \end{cases}$$

Por lo tanto,

$$P(X \geq 0) = \int_0^2 \frac{1}{5} dx = \frac{2}{5},$$

y

$$P(-5 \leq X \leq \frac{1}{2}) = \int_{-3}^{\frac{1}{2}} \frac{1}{5} dx = \frac{1}{5}\left(\frac{1}{2} + 3\right) = \frac{7}{10}.$$

Distribución normal

La distribución normal es quizás una de las más importantes, debido a que tiene un rol central tanto en la teoría de probabilidad como en la teoría estadística. Algunos autores la llaman distribución gaussiana en honor a Gauss, a quien se le considera el "padre" de esta distribución. La importancia de la distribución normal se debe al famoso teorema del límite central.

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,

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

Notación: $X \sim N(\mu, \sigma^2)$.

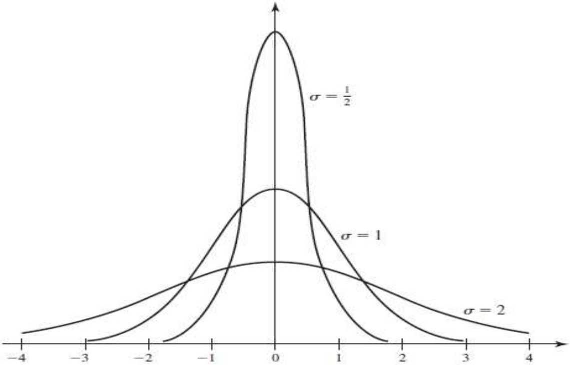

Figura 1: Gráfico de la fdp normal para $\mu = 0$ y $\sigma = 1/2, 1, 2$.

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,

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

Su fda, denota por $\Phi(\cdot)$, esta dada por,

$$\Phi(z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2}x^2\right\} dx, \;\; z \in \mathbb{R},$$

la cual está tabulada para varios valores de $z \in \mathbb{R}$; usualmente para $-4 < z < 4$.

Note que $\Phi(0) = 1/2$ y que $\Phi(z) + \Phi(-z) = 1, \forall\; z \in \mathbb{R}$.

Teorema 2.8

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

i) $\mathrm{E}(X) = \mu$

ii) $\mathrm{Var}(X) = \sigma^2$

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

Demostración 2.8

Calculemos la fgm de X. En efecto,

$$
\begin{aligned}
M(t) &= \frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty} \exp\left\{tx - \frac{1}{2\sigma^2}\,(x-\mu)^2\right\} dx \\
&= \exp\left\{\mu t + \frac{1}{2}\,\sigma^2 t^2\right\} \frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty} \exp\left\{-\frac{1}{2\sigma^2}\,\left(x-(\mu+\sigma^2 t)\right)^2\right\} dx \\
&= \exp\left\{\mu t + \frac{1}{2}\,\sigma^2 t^2\right\}, \;\; t \in \mathbb{R}.
\end{aligned}
$$

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

Demostración 2.9

Basta notar que,

$$\begin{aligned} M_Y(t) &= \exp\{at\}M_X(bt) = \exp\{at\}\exp\left\{\mu bt + \frac{}{2}\sigma^2 b^2 t^2\right\} \\ &= \exp\left\{(a + b\mu)t + \frac{1}{2}(\sigma b)^2 t^2\right\}, \end{aligned}$$

que corresponde a la fgm de una distribución normal con media $a + b\mu$ y varianza $\sigma^2 b^2$. Luego $Y \sim N(a + b\mu, b^2\sigma^2)$.

Ejemplo 2.11

Sea $X \sim N(1,4)$. Calcule, $P(0 \leq X < 1)$ y $P(X^2 > 4)$. Entonces,

$$
\begin{aligned}
P(0 \leq X < 1) &= P(-\frac{1}{2} \leq \frac{X-1}{2} < 0) \\
&= P(-\frac{1}{2} \leq Z < 0) \\
&= \Phi(0) - \Phi(-\frac{1}{2}) \\
&= 0.5 - 0.30854 \\
&= 0.19146.
\end{aligned}
$$

Similarmente,

$$\begin{aligned}
P\left(X^2 > 4\right) &= 1 - P(|X| \leq 2) \\
&= 1 - P(-2 \leq X \leq 2) \\
&= 1 - P(-\frac{3}{2} \leq \frac{X-1}{2} \leq \frac{1}{2}) \\
&= 1 - P(-\frac{3}{2} \leq Z \leq \frac{1}{2}) \\
&= 1 - \left\{\Phi(\frac{1}{2}) - \Phi(-\frac{3}{2})\right\} \\
&= 1 - \{0.69146 - 0.06681\} \\
&= 0.37535.
\end{aligned}$$

Ejemplo 2.12

Sean X_1 y X_2 las duraciones de dos dispositivos electrónicos. Asuma que $X_1 \sim N(40, 36)$ y $X_2 \sim N(45, 9)$. Si el dispositivo se va a utilizar durante 45 horas, qué dispositivo sería el preferido ? Si se va a utilizar durante 42 horas, cuál debería preferirse ?
Hay que encontrar qué dispositivo tiene una mayor probabilidad de vida útil de más de 45 horas:

$$\begin{aligned} P(X_1 > 45) &= P(Z > \frac{45 - 40}{6}) = 1 - \Phi(5/6) \\ &= 1 - 0.7995 = 0.2005 \end{aligned}$$

y

$$\begin{aligned} P(X_2 > 45) &= P(Z > \frac{45 - 45}{3}) = 1 - \Phi(0) \\ &= 1 - 0.5 = 0.5. \end{aligned}$$

Por lo tanto, se debe preferir el dispositivo X_2. Ahora, hay que encontrar qué dispositivo tiene una mayor probabilidad de vida útil de más de 42 horas. Cálculos similares producen,

$$\begin{aligned} P(X_1 > 42) &= 1 - \Phi(1/3) \\ &= 0.3707, \\ P(X_2 > 42) &= 1 - \Phi(-1) \\ &= 0.8413 \end{aligned}$$

En este caso también se debe preferir el dispositivo X_2.

Distribución Gama

La distribución gamma se usa de manera extensa en una variedad de áreas como, por ejemplo, para describir los intervalos de tiempo entre dos fallas consecutivas en el motor de un avión, o los intervalos de tiempo entre las llegadas de clientes a una cola en el cajero de un supermercado.

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 $\Gamma(n) = (n-1)!$, $n = 1, 2, \ldots$.

Notación: $X \sim Gama(\alpha, \lambda)$.

El orden de los parámetros es importante ya que α es un parámetro de forma, mientras que λ es un parámetro de escala.

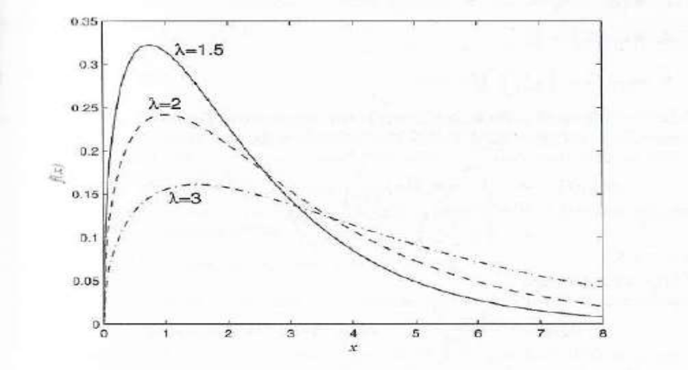

Figura 2: Gráfico de la fdp gama para $\alpha = 1.5$ y $\lambda = 1.5, 2, 3$.

Teorema 2.10

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

i) $\mathrm{E}(X) = \alpha/\lambda$

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

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

Demostración 2.10

Calculemos la fgm de X,

$$\begin{aligned} M(t) &= \int_0^\infty \exp(tx) \frac{\lambda}{\Gamma(r)} (\lambda x)^{\alpha-1} \exp(-\lambda x) dx \\ &= \left(\frac{\lambda}{\lambda - t}\right)^\alpha \int_0^\infty \frac{(\lambda - t)^\alpha}{\Gamma(\alpha)} x^{\alpha-1} \exp\{-(\lambda - t)x\} dx \\ &= \left(1 - \frac{t}{\lambda}\right)^{-\alpha}, \text{ si } \lambda - t > 0. \end{aligned}$$

Luego,

$$\mathrm{E}(X) = \frac{d}{dt}M(t)\Big|_{t=0} = \frac{\alpha}{\lambda} \text{ y } \mathrm{E}\left(X^2\right) = \frac{d^2}{dt^2}M(t)\Big|_{t=0} = \frac{\alpha^2+\alpha}{\lambda^2},$$

de donde sigue que $\mathrm{Var}(X) = \alpha/\lambda^2$.

Casos particulares:

i) Si $\alpha = 1 \Longrightarrow X \sim Exp(\lambda)$, distribución exponencial; en este caso

$$E(X) = \frac{1}{\lambda}, \quad \mathrm{Var}(X) = \frac{1}{\lambda^2}, \quad M(t) = (1-\frac{t}{\lambda})^{-1},\ t < \lambda$$

ii) $\alpha = \frac{\nu}{2}$ y $\lambda = \frac{1}{2} \Longrightarrow X \sim \chi^2_\nu$, distribución chi-cuadrado con ν grados de libertad; para este caso,

$$\mathrm{E}(X) = \nu, \quad \mathrm{Var}(X) = 2\nu, \quad M(t) = (1-2t)^{-\nu/2},\ t < 1/2.$$

iii) Si $\alpha > 1 \Longrightarrow X \sim Erlag(\alpha, \lambda)$, distribución de Erlang.

References

Blanco, L., Arunachalam, V. y Dharmaraja, S. (2012). *Introduction to Probability and Stochastic Processes.* John Wiley, New Jersey.

Casella, G. y Berger, R.L. (2002). *Statistical Inference.* Second Edition. Duxbury, California.