Formulario de Probabilidad

Roberto López

Conteo

Util para calcular probabilidades en espacios equiprobables(los eventos simples en el espacio tienen todos la misma probabilidad).

En ese caso la probabilidad es $P(A) = \frac{\text{número de casos favorables(donde ocurre } A)}{\text{número de casos totales.}}$.

Diagrama de árbol

Util para representar las posibles opciones o decisiones en cada etapa o paso, y asi saber cuantas formas totales hay, con la Regla de la multiplicación(multiplicar el número de opciones que hay para la decisión 1, por el número de opciones en la decisión 2, etc..).

Si tengo que elegir algo r veces, y cada vez hay las mismas n opciones(elección con reemplazo), el número total de formas es n^r.

Factorial. El número de formas de ordenar n objetos distintos es $n! = n(n-1)(n-2)\cdots(2)(1)$.

Permutaciones. De n objetos diferentes se eligen r objetos (sin reemplazo) y el orden importa($abc \neq bca$). El número de formas de hacer la elección es $nPr = \frac{n!}{(n-r)!}$

Combinaciones. De n objetos diferentes se eligen k objetos (sin reemplazo) y el orden NO importa($abc = bca = cba, etc..$). El número de formas de hacer la elección es $nCk = \binom{n}{k} = \frac{n!}{k!(n-k)!}$

Fórmulas básicas:

$P(A^c) = 1 - P(A)$

$P(A \cup B) = P(A) + P(B) - P(A \cap B)$

$P(A) = P(A \cap B) + P(A \cap B^C)$, entonces $P(A \cap B^C) = P(A) - P(A \cap B)$.

Leyes de DeMorgan:

$P(A^C \cap B^C) = P((A \cup B)^C) = 1 - P(A \cup B)$

$P(A^C \cup B^C) = P((A \cap B)^C) = 1 - P(A \cap B)$

Mas avanzadas:

$$P(A \cup B \cup C) = P(A) + P(B) + P(C) - [P(A \cap B) + P(A \cap C) + P(B \cap C)] + P(A \cap B \cap C)$$

$$P(A \cup B \cup C \cup D) = [P(A) + ... + P(D)] - [P(A \cap B) + P(A \cap C) + ... + P(B \cap D) + P(C \cap D)] + [P(A \cap B \cap C) + ... + P(B \cap C \cap D)] - P(A \cap B \cap C \cap D)$$

¿Como obtengo $P(A \cap B)$? Depende de la info. con la que cuento.

- $P(A \cap B) = P(A)P(B)$ cuando los eventos A y B sean independientes.
- $P(A \cap B) = P(A)P(B|A) = P(B)P(A|B)$
- $P(A \cap B) = P(A) + P(B) - P(A \cup B)$

A veces solo hay que hallar primero cual es $A \cap B$, ya después encontrar la probabilidad con conteo o sumando o integrando. A veces $P(A \cap B)$ es un área.
Ejemplo: Se tiran 2 dados. Resultan los números D_1, D_2.
$P(D_1 \leq 3 \cap D_1 + D_2 = 9) = P(D_1 = 3 \cap D_2 = 6) = P(D_1 = 3)P(D_2 = 6) = \frac{1}{36}$

Probabilidad condicional: $P(A|B) = \frac{P(A \cap B)}{P(B)}$

Teorema de Probabilidad Total: $P(A) = \Sigma P(A|B_i)P(B_i)$
útil cuando tengo la probabilidad del evento que quiero, pero solo la tengo ya habiendo ocurrido otros eventos. Por ejemplo: quiero $P(\text{llueve hoy})$ y tengo $P(llueve|Primavera), P(llueve|Verano), P(llueve|Otoño), P(llueve|Invierno)$.

Teorema de Bayes: $P(A|B) = \frac{P(B|A)P(A)}{\Sigma P(B|A_i)P(A_i)} \qquad \leftarrow \frac{P(A \cap B)}{P(B)}$
útil si me piden $P(A|B)$ y yo tengo $P(B|A)$ o esa es mas fácil de obtener.

Variables aleatorias
$f_X(x)$ es la función de masa de probabilidad o función de densidad
Si X es discreta entonces $f_X(x) = P(X = x)$.

$F_x(x) = P(X \leq x)$ es la f. de probabilidad acumulada o f. de distribución.

Valor esperado(o media o esperanza): $\mu = E[X] = \int_{-\infty}^{\infty} xf(x)dx$, o con Σ.
Ley del estadístico inconsciente: $E[g(X)] = \int_{-\infty}^{\infty} g(x)f(x)dx$
Por ejemplo $E[X^2] = \int_{-\infty}^{\infty} x^2 f(x)dx$

La Varianza es $Var(X) = E[X^2] - E[X]^2$. Se denota $\sigma^2 = Var(X)$.

Propiedades de $E[X]$ y $Var(X)$
$E[c] = c, \qquad E[cX] = cE[X], \qquad E[X + Y] = E[X] + E[Y]$
$Var(c) = 0, \qquad Var(cX) = c^2 Var(X), \qquad Var(X + c) = Var(X),$
$Var(X \pm Y) = Var(X) + Var(Y) \pm 2cov(X, y)$

Distribuciones principales
Discretas
$X \sim Bernoulli(p)$ si $f(x) = p^x(1-p)^{1-x} I_{\{0,1\}}(x)$
En ese caso $E[X] = p, \quad V(X) = p(1-p)$

$X \sim Binomial(n, p)$ si $f(x) = \binom{n}{x} p^x (1-p)^{n-x} I_{\{0,1,\ldots,n\}}(x)$
$E[X] = np, \quad V(X) = np(1-p), \quad X =$ número de éxitos en n ensayos.

$X \sim Poisson(\lambda)$ si $f(x) = \frac{e^{-\lambda}\lambda^x}{x!} I_{\{0,1,\ldots\}}(x)$
En ese caso $E[X] = \lambda, \quad V(X) = \lambda$
Esta aplica cuando $X =$número de eventos ocurridos en cierto tiempo. O tal vez número de objetos que hay en cierto espacio(longitud, área, volumen).

$X \sim Geométrica(p)$ si $f(x) = (1-p)^{x-1} p I_{\{1,2,\ldots\}}(x)$
Esta aplica cuando $X =$ número de ensayos hasta lograr el primer éxito.
En este caso $E[X] = \frac{1}{p}, \quad Var(X) = \frac{1-p}{p^2}, \quad F_X(x) = 1-(1-p)^x$

$X \sim Geométrica(p)$ si $f(x) = (1-p)^x p I_{\{0,1,2,\ldots\}}(x)$
Esta aplica cuando $X =$ número de fracasos hasta lograr el primer éxito.
En este caso $E[X] = \frac{1}{p} - 1 = \frac{1-p}{p}, \quad Var(X) = \frac{1-p}{p^2}, F_X(x) = 1-(1-p)^{x+1}$

$X \sim BinomialNegativa(r,p)$ si $f(x) = \binom{x-1}{r-1} p^r (1-p)^{x-r} I_{\{r,r+1,\ldots\}}(x)$
Esta aplica cuando $X =$número de experimentos hasta lograr el r-ésimo éxito.
En este caso $E[X] = \frac{r}{p}, \quad Var(X) = \frac{r(1-p)}{p^2}$.
También se puede $X \sim BinNeg(r,p)$ si $f(x) = \binom{r+x-1}{r-1} p^r (1-p)^x I_{\{0,1,2,\ldots\}}(x)$
Esta aplica cuando $X =$número de fracasos hasta lograr el r-ésimo éxito.
En este caso $E[X] = r(\frac{1}{p} - 1), \quad Var(X) = \frac{r(1-p)}{p^2}$.

$X \sim HiperGeométrica(A,B,N,n)$ si $f(x) = \frac{\binom{A}{x}\binom{B}{n-x}}{\binom{N}{n}}$
En este caso $E[X] = n\frac{A}{N}, \quad Var(X) = n\frac{A}{N}(\frac{N-A}{N})(\frac{N-n}{N-1})$.
Esta aplica cuando hay N objetos, de los cuales hay 2 tipos, y hay A del tipo 1 y B del tipo 2 ($A+B=N$). Se elegirán n objetos de los N.
$X =$número de objetos elegidos que pertenecen al tipo 1.

Continuas
$X \sim Uniforme(a,b)$ si $f(x) = \frac{1}{b-a} I_{(a,b)}(x)$.
En ese caso $E[X] = \frac{a+b}{2}, \quad Var(X) = \frac{(b-a)^2}{12}, \quad F_X(x) = \frac{x-a}{b-a}$.

$X \sim Exp(\lambda)$ si $f(x) = \lambda e^{-\lambda x} I_{(0,\infty)}(x)$.
En ese caso $E[X] = \frac{1}{\lambda}, \quad Var(X) = \frac{1}{\lambda^2}, \quad F_X(x) = 1 - e^{-\lambda x}$.
Si defino $\beta = \frac{1}{\lambda}$ entonces se puede $X \sim Exp(\beta)$ si $f(x) = \frac{1}{\beta} e^{-\frac{1}{\beta}x} I_{(0,\infty)}(x)$.
Ahora $E[X] = \beta, \quad Var(X) = \beta^2, \quad F_X(x) = 1 - e^{-\frac{1}{\beta}x}$.

$X \sim Gamma(\alpha,\beta)$ si $f(x) = \frac{x^{\alpha-1}e^{-x/\beta}}{\Gamma(\alpha)\beta^\alpha} I_{(0,\infty)}(x)$.
En ese caso $E[X] = \alpha\beta, \quad Var(X) = \alpha\beta^2$
También se puede usar $\lambda = 1/\beta$ y queda $f(x) = \frac{x^{\alpha-1}e^{-\lambda x}}{\Gamma(\alpha)} \lambda^\alpha I_{(0,\infty)}(x)$.
Ahora $E[X] = \alpha/\lambda, Var(X) = \alpha/\lambda^2$
En ambas se usa la función Gamma Γ.
Se puede necesitar esto: $\Gamma(t+1) = t\Gamma(t), \quad \Gamma(n+1) = n!, \quad \Gamma(\frac{1}{2}) = \sqrt{\pi}$.

Casos particulares de la Gamma:
Si $\alpha = 1$ entonces $X \sim Gamma(\alpha = 1, \beta)$ se vuelve $X \sim Exp(\beta)$.
Por otro lado, si $\beta = 2$ (o $\lambda = \frac{1}{2}$) entonces $X \sim Gamma(\alpha, \beta = 2) = \chi^2_{2\alpha}$, que es la distribución Ji cuadrada.

$X \sim Normal(\mu, \sigma^2)$ si $f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} I_{(-\infty,\infty)}(x)$
En este caso $E[X] = \mu, \quad Var(X) = \sigma^2$.
$Z \sim$Normal estándar si $Z \sim N(0,1)$. (tiene media 0 y varianza 1)
Teorema: Si $X \sim N(\mu, \sigma^2)$ y defino $Z = \frac{X-\mu}{\sigma}$, entonces $Z \sim N(0,1)$.
Eso se conoce como estandarizar.

Truco para calcular $E[X^2]$
Cuando hablamos de alguna de las distrib. principales y debemos encontrar $E[X^2]$, en lugar de integrar podemos despejarlo de la fórmula de la varianza: $E[X^2] = Var(X) + E[X]^2$.

Transformaciones
Tenemos una v.a. X con su $f_X(x)$ y/o su $F_X(x)$.
Nos dicen que $Y = g(X)$, por ejemplo $Y = X^3$.
Se pide hallar $f_Y(y)$. Hay 2 maneras:
I) Método de la F: $F_Y(y) = P(Y \leq y) = P(g(X) \leq y) =$ alguna probabilidad con X(despejar X, resolver la desigualdad). Queda en términos de $F_X()$.
Sustituir la F_X si se tiene, o derivar ambos lados, para ya tener $f_Y(y)$.
II) Teorema de cambio de variable. Seguir estos pasos:
1) Despejar X : queda $x = g^{-1}(y) =$algo. 2) Derivar: $\frac{dx}{dy} =$algo.
3) Sustituir en esta fórmula: $f_Y(y) = f_X(x \text{ despejada }) |\frac{dx}{dy}|$.

Momentos de la v.a. X: $E[X] = \mu, \quad E[X^2], \quad E[X^3]$, ...
Momentos centrales: $E[X - \mu] = 0, E[(X-\mu)^2] = \sigma^2, E[(X-\mu)^3]$, etc...
Función generadora de momentos es $M_X(t) = E[e^{tX}] = \int e^{tx} f(x)dx$ o con Σ.
Teorema: $E[X^k] = \frac{d^k M_X(t)}{dt^k}(0)$. (Es derivar k veces la f.g.m. y evaluar en $t = 0$).

Si $X \sim Bernoulli(p)$, entonces $M_X(t) = (1 - p + pe^t)$.
Si $X \sim Binomial(n,p)$, entonces $M_X(t) = (1 - p + pe^t)^n$.
Si $X \sim Poisson(p)$, entonces $M_X(t) = e^{\lambda(e^t - 1)}$.
Si $X \sim Uniforme(a,b)$, entonces $M_X(t) = \frac{e^{bt} - e^{at}}{t(b-a)}$.
Si $X \sim Exponencial(\lambda)$, entonces $M_X(t) = \frac{\lambda}{\lambda - t} = \frac{1}{1 - \beta t}$.
Si $X \sim Normal(\mu, \sigma^2)$, entonces $M_X(t) = e^{\mu t + \frac{1}{2}\sigma^2 t^2}$.
Si $X \sim Gamma(\alpha, \beta)$, entonces $M_X(t) = (1 - \beta t)^{-\alpha}$.

Vectores aleatorios
$f(x,y)$ es la f. de prob. conjunta.
En las v.a. discretas: $f(x,y) = P(X = x \cap Y = y)$.

Las f. de prob. marginales son $f_X(x)$ y $f_Y(y)$, y se obtienen asi:
$f_X(x) = \int_{S_Y} f(x,y)dy$ y $f_Y(y) = \int_{S_X} f(x,y)dx$ donde S_X, S_Y son los soportes.
En las discretas hay que cambiar $\int$ por Σ.

Las f. de prob. condicionales son $f(x|y) = \frac{f(x,y)}{f_Y(y)}$ y $f(y|x) = \frac{f(x,y)}{f_X(x)}$.
Luego puedo calcular $E[X|Y] = \int xf(x|y)dx$ y $E[X^2|Y] = \int x^2 f(x|y)dx$ para luego tener $Var(X|Y) = E[X^2|Y] - E[X|Y]^2$.

Teorema: $E[Y] = E[E[Y|X]]$ y además $Var(Y) = Var(E[Y|X]) + E[Var(Y|X)]$.

$E[g(X,Y)] = \int\int g(x,y)f(x,y)dxdy$.
Por ejemplo $E[XY] = \int\int xyf(x,y)dxdy$, que se usa para calcular la covarianza:
$\sigma_{XY} = cov(x,y) = E[XY] - E[X]E[Y]$.
La correlación es $\rho = \frac{cov(x,y)}{\sqrt{Var(X)}\sqrt{Var(Y)}}$.

Propiedades de cov
$cov(X,Y) = cov(Y,X)$, $cov(a,Y) = cov(X,b) = 0$ donde a, b son constantes.
$cov(X,X) = Var(X)$
$cov(X+Y,Z) = cov(X,Z) + cov(Y,Z)$

Transformaciones bivariadas o multivariadas
Tengo $f_{X,Y}(x,y)$. Me dicen que $U = g(X,Y)$ y $V = h(X,Y)$.
Piden encontrar la f. de prob. conjunta $f_{U,V}(u,v)$.
Paso 1: despejar X, Y en términos de U, V.
Paso 2: derivar eso. Queda la matriz Jacobiana $J = \begin{pmatrix} \frac{\partial X}{\partial U} & \frac{\partial X}{\partial V} \\ \frac{\partial Y}{\partial U} & \frac{\partial Y}{\partial V} \end{pmatrix}$.
Paso 3: Sustituir en esta fórmula: $f_{U,V}(u,v) = f_{X,Y}(x \text{ despejada}, y \text{ despejada})|det(J)|$.
Paso 4: incluir el soporte del vector (U,V).

Func. gen. de momentos conjunta: $M_{X,Y}(t,r) = E[e^{tx+ry}] = \int\int e^{tx+ry} f(x,y)dxdy$
Propiedades
$M_{X,Y}(t,0) = M_X(t)$, $M_{X,Y}(0,r) = M_Y(r)$
$E[X^a Y^b] = \frac{\partial^{a+b} M}{\partial t^a \partial r^b}|_{(0,0)}$

Distribución Normal Bivariada. $\begin{pmatrix} X \\ Y \end{pmatrix} \sim N_2\Big(\mu = \begin{pmatrix} \mu_x \\ \mu_y \end{pmatrix}, \Sigma = \begin{pmatrix} \sigma_x^2 & \sigma_{x,y} \\ \sigma_{x,y} & \sigma_y^2 \end{pmatrix}\Big)$
si $f(x,y) = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}} e^{-\frac{1}{2(1-\rho^2)}[(\frac{X-\mu_x}{\sigma_x})^2 - 2\rho(\frac{X-\mu_x}{\sigma_x})(\frac{Y-\mu_y}{\sigma_y}) + (\frac{Y-\mu_y}{\sigma_y})^2]}$.
Las marginales son $X \sim N(\mu_x, \sigma_x^2)$ y $Y \sim N(\mu_y, \sigma_y^2)$.
Ademas se tiene $(X|Y=y) \sim N(\mu_x + \rho\frac{\sigma_x}{\sigma_y}(y-\mu_y), \sigma_x^2(1-\rho^2))$.
La f.g.m del vector es $M_{X,Y}(t,r) = e^{\mu_x t + \mu_y r + \frac{1}{2}[\sigma_x^2 t^2 + 2\rho\sigma_x\sigma_y tr + \sigma_y^2 r^2]}$.
Teorema: Si $\begin{pmatrix} U \\ V \end{pmatrix} = A \begin{pmatrix} X \\ Y \end{pmatrix} + b$, entonces $\begin{pmatrix} U \\ V \end{pmatrix} \sim N_2(A\mu + b, A\Sigma A^T)$.