Variáveis Aleatórias e Distribuições de Probabilidade

Variáveis aleatórias

Variável Aleatória: Uma variável aleatória utiliza-se para expressar os resultados de uma experiência aleatória.

Espaço Amostral: Conjunto de todos os resultados possíveis de uma experiência aleatória.

Representação de populações através das funções de probabilidade

Seja Y uma ***variável aleatória discreta***:

A função de probabilidade da variável Y, que se denota por $p(y)$, é a função que associa a cada valor particular que a variável pode tomar (y) a probabilidade de Y ser igual a y:

$$p(y) = Probabilidade\,(Y = y).$$

A função de distribuição (ou função de probabilidade acumulada) da variável Y, que se denota por $F(y)$, é a função que associa a cada valor particular que a variável pode tomar (y) a probabilidade de Y ser menor ou igual a y:

$$F(y) = Probabilidade\,(Y \leq y).$$

Seja X uma ***variável aleatória contínua***:

A probabilidade de a variável X tomar um valor situado dentro de um intervalo finito $[a, b]$ obtém-se a partir da expressão seguinte:

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

A função de distribuição (ou função de probabilidade acumulada) da variável X, $F(x)$, associa a cada valor particular x a probabilidade de X ser menor ou igual a x. Em notação simbólica, a função de distribuição de X é definida pela expressão:

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(u)\,du.$$

É de notar que no integral que figura na expressão o limite superior de integração é x, pelo que há que considerar uma variável de integração distinta de x, no caso a variável u.

As funções de probabilidade podem ser encaradas como formas de representação de populações.

Tal como as distribuições de frequências, calculadas com base nos dados amostrais, são susceptíveis de serem caracterizadas por estatísticas, também as *funções de probabilidade* associadas às *populações* podem ser caracterizadas por um conjunto de parâmetros.

Para uma dada população, as funções de probabilidade e os parâmetros que lhes estão associados são fixos, em contraste com as estatísticas relativas às frequências observadas numa amostra, que variam de amostra para amostra.

Parâmetros das funções de probabilidade de variáveis aleatórias

Parâmetros de localização:

	Variáveis Discretas	Variáveis Contínuas
Valor esperado (μ)	$\mu_Y = E(Y) = \sum_y y \cdot p(y)$	$\mu_X = E(X) = \int_{-\infty}^{+\infty} x \cdot f_X(x)dx$
Mediana (η)	Seja y_0 o valor mínimo que uma variável discreta Y pode tomar e $F(y)$ a função de probabilidade acumulada. Se $F\,(y0) \geq 0.5$, a mediana é igual a y_0, ou seja $\eta_y = y_0$; Se $F(y_0 < 0.5)$, então $\eta_y = y_1$, onde y_1 denota o valor particular de Y que satisfaz as seguintes condições: $F(y_1) \geq 0.5\ e\ F(y_1^-) < 0.5$ (y_1^- representa o valor particular de Y imediatamente inferior a y_1).	Para uma variável X contínua, se a equação $F(x) = 0.5$ tem: - uma solução, x corresponde à mediana, η_x; - mais do que uma solução, a mediana é definida por: $\eta_x = \frac{X_{min} + X_{max}}{2}$ (X_{min} e X_{max} representam o mínimo e o máximo do conjunto de soluções, respectivamente).
Moda (ζ)	A moda é o valor para o qual a função de probabilidade toma o valor máximo.	A moda é o valor para o qual a função densidade de probabilidade toma o valor máximo.

Parâmetros de dispersão:

	Variáveis Discretas	Variáveis Contínuas
Desvio absoluto médio (δ)	$\delta_Y = \sum_y \lvert y - \mu_y \rvert \cdot p(y)$	$\delta_X = \int_{-\infty}^{+\infty} \lvert x - \mu_x \rvert \cdot f_X(x)dx$
Variância (σ^2 ou *Var()*) $\sigma_X^2 = Var(X) = E[(X - \mu_X)^2\,] = E(X^2) - E^2(X)$	$\sigma_Y^2 = Var(Y) = \sum_y (y - \mu_Y)^2 \cdot p(y)$	$\sigma_X^2 = Var(X) = \int_{-\infty}^{+\infty} (x - \mu_x)^2 \cdot f_X(x)dx$
Desvio padrão (σ)	$\sigma_Y = \sqrt{\sigma_Y^2}$	$\sigma_X = \sqrt{\sigma_X^2}$

Outros parâmetros:

	Variáveis Discretas	Variáveis Contínuas
Momento ordinário de ordem i (μ_i') ($i = 1, 2, \ldots$)	$\mu_i' = \sum_y y^i \cdot p(y)$	$\mu_i' = \int_{-\infty}^{+\infty} x^i \cdot f_X(x)dx$
Momento centrado de ordem *i* (μ_i) ($i = 1, 2, \ldots$)	$\mu_i = \sum_y (y - \mu_Y)^i \cdot p(y)$	$\mu_i = \int_{-\infty}^{+\infty} (x - \mu_x)^i \cdot f_X(x)dx$
Coeficiente populacional de assimetria (γ_1)	$\gamma_1 = \frac{\mu_3}{\sigma^3}$	
Coeficiente Populacional de Kurtose (γ_2)	$\gamma_2 = \frac{\mu_4}{\sigma^4} - 3$	

Definição de distribuições conjuntas, marginais e condicionais

No estudo de experiências aleatórias é muitas vezes necessário caracterizar mais do que uma variável aleatória a partir do mesmo espaço amostral. Nestas circunstâncias é importante pôr em evidência não só as propriedades individuais das variáveis mas também as propriedades conjuntas (i.e., aquelas que dizem respeito à forma como tais variáveis se encontram associadas). Isto envolve a definição de *distribuições conjuntas*, de *distribuições marginais* e de *distribuições condicionais*. No quadro seguinte apresentam-se as expressões relativas à função conjunta de probabilidade (ou de densidade de probabilidade), à função marginal de probabilidade (ou de densidade de probabilidade) e às funções condicionais de probabilidade (ou de densidade de probabilidade).

	Variáveis Discretas	Variáveis Contínuas
Distribuição conjunta	$p_{H,W}(h,w) = P(H = h, W = w)$	$Prob\{(X,Y) \in D\} = \iint_D f_{X,Y}(x,y)\,dx\,dy$
Distribuição marginal	$p_H(h) = \sum_w p_{H,W}(h,w)$ $p_W(w) = \sum_h p_{H,W}(h,w)$	$f_X(x) = \int_R f_{X,Y}(x,y)\,dy$ $f_Y(y) = \int_R f_{X,Y}(x,y)\,dx$
Distribuição condicional	$p_{H\|W}(h\|w) = \frac{p_{H,W}(h,w)}{p_W(w)} = \frac{p_{H,W}(h,w)}{\sum_h p_{H,W}(h,w)}$ $p_{W\|H}(w\|h) = \frac{p_{H,W}(h,w)}{p_H(h)} = \frac{p_{H,W}(h,w)}{\sum_w p_{H,W}(h,w)}$	$f_{X\|Y}(x\|y) = \frac{f_{X,Y}(x,y)}{f_Y(y)} = \frac{f_{X,Y}(x,y)}{\int_R f_{X,Y}(x,y)\,dx}$ $f_{Y\|X}(y\|x) = \frac{f_{X,Y}(x,y)}{f_X(x)} = \frac{f_{X,Y}(x,y)}{\int_R f_{X,Y}(x,y)\,dy}$

Note-se que estas expressões podem ser facilmente adaptadas para o caso em que uma variável é discreta e outra é contínua.

Independência entre variáveis

A independência entre duas variáveis implica que entre elas não existe qualquer relação, isto é, qualquer que seja o valor particular que uma delas tome não se altera a distribuição de probabilidade da outra. Daqui resulta a seguinte condição necessária e suficiente de independência entre duas variáveis aleatórias discretas H e W e entre duas variáveis contínuas X e Y:

$$\forall h, w: p_{H,W}(h,w) = p_H(h) \cdot p_W(w) \qquad \forall x, y: f_{X,Y}(x,y) = f_X(x) \cdot f_Y(y)$$

Se esta condição for satisfeita para todos os valores particulares das variáveis, então estas serão efectivamente independentes. Inversamente, se para algum par de valores a relação não for satisfeita, as variáveis não são independentes.

Covariância e correlação

A covariância e a correlação são medidas do grau de relacionamento linear entre duas variáveis.
A covariância das variáveis aleatórias discretas H e W, que se denota por $\gamma_{H,W}$ ou $Cov(h,w)$, é definida por:

$$\gamma_{H,W} = \sum_H \sum_W (h - \mu_H) \cdot (w - \mu_W) \cdot p_{H,W}(h,w)$$

A covariância de X e Y, variáveis contínuas, é dada por:

$$\gamma_{X,Y} = \iint_{R^2} (x - \mu_X) \cdot (y - \mu_Y) \cdot f_{X,Y}(x,y) \cdot \mathrm{d}x \cdot \mathrm{d}y$$

O coeficiente de correlação (ou, simplesmente, a correlação) é obtido dividindo a covariância pelo produto dos desvios padrão das variáveis em causa.

$$\rho_{H,W} = \frac{\gamma_{H,W}}{\sigma_H \cdot \sigma_W}$$

O coeficiente de correlação pode tomar valores situados entre –1 e 1 (isto é, $-1 \leq \rho_{H,H} \leq 1$).

$\rho_{H,W} = 1$ indica uma relação linear perfeita entre as variáveis consideradas, de tal forma que quando H aumenta W também aumenta (isto é, $W = a + b \cdot H$ com $b > 0$)

$\rho_{H,W} = -1$ indica uma relação linear perfeita entre as variáveis consideradas, de tal forma que quando uma aumenta a outra diminui (isto é, $W = a + b \cdot H$ com $b < 0$).

$\rho_{H,W} = 0$ indica que o grau de relacionamento linear entre as variáveis é nulo.

É evidente que se as duas variáveis aleatórias forem independentes então não existirá qualquer relação entre elas, sendo portanto nulo o grau de relacionamento linear (o que implica que $\rho_{H,W} = 0$).

No entanto, se o coeficiente de correlação for nulo não se pode daí depreender que as variáveis são independentes (uma vez que podem existir relações não-lineares entre elas).

Variáveis Transformadas:

Distribuições de variáveis transformadas

V***ariáveis aleatórias discretas***: Por simples inspecção e enumeração.

V***ariáveis aleatórias contínuas***: Conhecidas a função densidade de probabilidade da v.a. X ($f_X(x)$) e a função de transformação ($W = \Phi(X)$), a função de densidade de probabilidade da v.a. W ($f_W(w)$) por:

$$\mathrm{f_W(w)} = \sum_i \frac{f_X(x)}{\left|\frac{dw}{dx}\right|}\Bigg|_{\substack{\mathrm{X}=\Phi^{-1}(W) \\ x \in \text{região } i \\ w \in \text{região } i}}, (i = 1, 2, \ldots, K).$$

Onde o domínio de X é dividido em K regiões de forma a que a função $\Phi(X)$), seja monótona e, logo, invertível em cada região.

Valor esperado e variância de variáveis transformadas

Transformações lineares de variáveis

Considere-se uma variável aleatória qualquer Y, discreta ou contínua, e admita-se que os seus valores particulares são transformados por aplicação de uma função linear $f(Y) = a + b \cdot Y$, onde a e b são constantes. Como resultado desta operação, fica definida uma nova variável aleatória (X):

$$X = a + b \cdot Y$$

O valor esperado e a variância da variável transformada (X) podem ser calculados directamente a partir dos parâmetros da variável original (Y):

$$\mu_X = \mathrm{E(X)} = \mathrm{E(a + b \cdot Y)} = \mathrm{a + b \cdot E\,(Y)} \qquad \sigma_X^2 = \mathrm{Var(X)} = \mathrm{Var(a + b \cdot Y)} = \mathrm{b^2 \cdot Var(Y)}$$

Transformações não lineares de variáveis

Considere-se uma variável aleatória qualquer Z, discreta ou contínua, e admita-se que os seus valores particulares são transformados por aplicação de uma função não linear ϕ(Z). Como resultado desta operação, fica definida uma nova variável aleatória $W = \Phi(Z)$.

Quando **não se conhece** a função (de densidade) de probabilidade de Z, o <u>valor esperado</u> e a <u>variância</u> da variável transformada (W) podem ser calculados, de forma **aproximada**, a partir das expressões:

$$\mathrm{E(W)} = \mathrm{E}[\Phi(\mathrm{Z})] \approx \Phi[\mathrm{E(Z)}]$$

$$\mathrm{Var(W)} = \mathrm{Var}[\Phi(\mathrm{Z})] \approx \left[\frac{\mathrm{d}\Phi}{dZ}\right]^2_{\mathrm{Z}=\mu_\mathrm{Z}} \cdot \mathrm{Var(Z)}$$

Quando **se conhece** a função (de densidade) de probabilidade de Z, o <u>valor esperado</u> e a <u>variância</u> da variável transformada (W) podem ser calculados, de forma **exacta**, a partir das expressões

$$\mathrm{E(W)} = E[\Phi(Z)] = \textstyle\sum_Z \Phi(z) \cdot p_Z(z) \qquad Var(W) = Var[\Phi(Z)] = \sum_Z (\Phi(z) - E[\Phi(Z)])^2 \cdot p_Z(z)$$

$$\mathrm{E(W)} = E[\Phi(Z)] = \textstyle\int_{-\infty}^{+\infty} \Phi(z) \cdot f_Z(z) \qquad Var\mathrm{(W)} = Var[\Phi(Z)] = \int_{-\infty}^{+\infty} (\Phi(z) - E[\Phi(Z)])^2 \cdot f_Z(z)$$

Note-se que, ao contrário da situação anterior, neste caso conhece-se a função de probabilidade de Z ($p_Z(z)$), no caso discreto, ou a função densidade de probabilidade de Z ($f_Z(z)$), no caso contínuo.

Distribuições de pares de variáveis transformadas

V***ariáveis aleatórias discretas***: Por simples inspecção e enumeração.

V***ariáveis aleatórias contínuas***: Conhecidas a função densidade de probabilidade conjunta das variáveis X e Y ($f_{XY}(x,y)$) e as funções de transformação ($W = \Phi_1(X,Y)$ e $V = \Phi_2(X,Y)$), a função de densidade de probabilidade conjunta das variáveis W e V ($f_{W,V}(w,z)$) por:

$$\mathrm{f_{W,Z}(w,z)} = \sum_i [f_{XY}(x,y) \cdot |J(z,w)|]\Big|_{\substack{\mathrm{X}=\mathrm{G_1}(w,z) \\ \mathrm{y}=\mathrm{G_2}(w,z) \\ x,y \in \text{região } i \\ w,z \in \text{região } i}}, (i = 1, 2, \ldots, K).$$

$$J(z,w) = \frac{d(x,y)}{d(w,z)} = \begin{vmatrix} \frac{dx}{dw} & \frac{dx}{dz} \\ \frac{dy}{dw} & \frac{dy}{dz} \end{vmatrix}$$

Onde o domínio de X, Y é dividido em K regões de modo a que as funções $\Phi_1(X,Y)$ e $\Phi_2(X,Y)$ possam ser resolvidas univocamente para X e Y em função de Z e W, isto é existe a transformação inversa $X = G_1(W,Z)$ e $Y = G_2(W,Z)$.

É ainda necessário que:

A transformação e a respectiva inversa sejam contínuas;

As derivadas parciais dx/dw, dx/dz, dy/dw e dy/dz, existam e sejam contínuas;

O Jacobiano, $J(z,w)$, da transformação inversa seja diferente de zero.

Valor esperado e variância de variáveis transformadas

Combinação linear de variáveis

Considerem-se as variáveis aleatórias H e W, discretas ou contínuas, e admita-se que os seus valores particulares são transformados por aplicação de uma função linear $\mathrm{f(H,W) = a + b \cdot H + c \cdot W}$, onde a, b e c são constantes. Como resultado desta operação fica definida uma nova variável aleatória V (em que $\mathrm{V = f(H,W) = a + b \cdot H + c \cdot W}$).
O <u>valor esperado</u> e a <u>variância</u> de V podem ser calculados directamente, de forma **exacta**, a partir dos parâmetros das variáveis originais (H e W):

$$E(V) = E(\ a + b \cdot H + c \cdot W) = a + b \cdot E(H) + c \cdot E(W)$$

$$Var(V) = Var(\ a + b \cdot H + c \cdot W) = b^2 \cdot Var(H) + 2 \cdot b \cdot c \cdot Cov(H,W) + c^2 \cdot Var(W)$$

Combinação não linear de variáveis

Considere-se o caso em que a transformação: $\mathrm{V = \Phi(H,W)}$ é não linear.

Quando **não se conhece** a função (de densidade) de probabilidade conjunta de H e W, o <u>valor esperado</u> e a <u>variância</u> da variável transformada (V) podem ser calculados, de forma **aproximada**, a partir das expressões:

$$\mathrm{E(V) = E[\Phi(H,W))] \approx \Phi[E(H),E(W)]}$$

$$\mathrm{Var(V) = Var[\Phi(H,W)]} \approx \left[\frac{\mathrm{d\Phi}}{dH}\right]^2_{\substack{\mathrm{H=\mu_H} \\ W=\mu_W}} \cdot \mathrm{Var(H)} + 2 \cdot \left[\frac{\mathrm{d\Phi}}{dH}\right]_{\substack{\mathrm{H=\mu_H} \\ W=\mu_W}} \cdot \left[\frac{\mathrm{d\Phi}}{dW}\right]_{\substack{\mathrm{H=\mu_H} \\ W=\mu_W}} \cdot \mathrm{Cov(H,W)} + \left[\frac{\mathrm{d\Phi}}{dW}\right]^2_{\substack{\mathrm{H=\mu_H} \\ W=\mu_W}} \cdot \mathrm{Var(W)}$$

Quando **se conhece** a função (de densidade) de probabilidade conjunta de H e W (X e Y), o <u>valor esperado</u> e a <u>variância</u> da variável transformada (V) podem ser calculados, de forma **exacta**, a partir das expressões:

$$\mathrm{E(V)} = E[\Phi(H,W)] = \sum_H \sum_W \Phi(H,W) \cdot p_{H,W}(h,w) \qquad Var(V) = Var[\Phi(H,W)] = \sum_H \sum_W (\Phi(H,W) - E[\Phi(H,W)])^2 \cdot p_{H,W}(h,w)$$

$$\mathrm{E(V)} = E[\Phi(X,Y)] = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \Phi(X,Y) \cdot f_{X.Y}(x,y) \qquad Var(\mathrm{V}) = Var[\Phi(X,Y)] = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} (\Phi(X,Y) - E[\Phi(X,Y)])^2 \cdot f_{X,Y}(x,y)$$

Note-se que, ao contrário da situação anterior, neste caso conhece-se a função de probabilidade conjunta de H e W ($p_{H,W}(h,w)$), no caso discreto, ou a função densidade de probabilidade conjunta de X e Y ($f_{X,Y}(x,y)$), no caso contínuo.

Covariância definida como valor esperado de uma variável transformada

Seja a v. a. $V = (H - \mu_H) \cdot (W - \mu_W)$, o valor esperado de V é dado do

$$E(V) = E[(X - \mu_X) \cdot (Y - \mu_Y)] = E(XY) - E(X) \cdot E(Y) = \gamma_{H,W}$$

Ou seja, $$\gamma_{H,W} = Cov(X,Y) = E[(X - \mu_X) \cdot (Y - \mu_Y)] = E(XY) - E(X) \cdot E(Y)$$

Formulário adaptado de:
Estatística
Rui Campos Guimarães, José A. Sarsfield Cabral
Verlag Dashöfer