Erasmus School of Economics

MOOC Econometrics

Lecture P.1 on Building Blocks: Random Variables

Erik Kole

Erasmus University Rotterdam

Density Functions

Let x be a random variable.

$$P[a \leq x \leq b] = \int_a^b f(v)dv$$

- f: probability density function (pdf)
- for all v: $f(v) \geq 0$, $\int_{-\infty}^{\infty} f(v) = 1$

$$P[x \leq b] = \int_{-\infty}^{b} f(v)dv = F(b)$$

- F: cumulative density function (cdf)
- $\lim_{v \to -\infty} F(v) = 0$ and $\lim_{v \to \infty} F(v) = 1$.

Expectation: mean and variance

Expectation operator E:

mean: $\mu = E[x] = \int v \cdot f(v)dv$

general: $E\left[g(x)\right] = \int g(v)f(v)dv \neq g(\mu)$

variance: $\sigma^2 = \text{var}[x] = E\left[(x-\mu)^2\right] = \int (v-\mu)^2 f(v)dv$

- for all v, $(v-\mu)^2 \geq 0$ and $f(v) \geq 0$, so $\sigma^2 \geq 0$.
- standard deviation $\sigma = \sqrt{\text{var}[x]}$.

Expectation of linear functions

Let x be a random variable, $E[x] = \mu_x$, $\text{var}[x] = \sigma_x^2$.

$$y = ax + b, \quad a, b \text{ constant}$$

$$\begin{aligned} E[y] = E[ax+b] &= \int (av+b)f(v)dv \\ &= \int a \cdot v \cdot f(v)dv + \int b \cdot f(v)dv \\ &= a \int v \cdot f(v)dv + b \int f(v)dv \\ &= aE[x] + b \cdot 1 = a\mu_x + b \end{aligned}$$

Variance of linear function

Test

Let x be a random variable, $E[x] = \mu_x$, $\text{var}[x] = \sigma_x^2$. Consider the function $y = ax + b$. What is the variance of y?

Answer

$$\begin{aligned}
\text{var}[y] &= E\left[(y-\mu_y)^2\right] = E\left[\left((ax+b) - (a\mu_x + b)\right)^2\right] \\
&= E\left[(ax - a\mu_x)^2\right] \\
&= E\left[a^2(x-\mu_x)^2\right] \\
&= a^2 E\left[(x-\mu_x)^2\right] \\
&= a^2\sigma_x^2
\end{aligned}$$

Two random variables

Let x and y be random variables.

$$P[a \le x \le b, c \le y \le d] = \int_a^b \left(\int_c^d f(v,w)\, dw \right) dv$$

- $f(v,w)$: joint pdf.
- for all v, w: $f(v,w) \ge 0$; $\iint f(v,w)\, dw\, dv = 1$.

Marginal density:

$$f_x(v) = \int f(v,w) dw$$

Mean, variance and covariance

Mean and variance: use marginal density

$$\mu_x = E[x] = \int v\, f_x(v)\, dv = \iint v\, f(v,w)\, dw\, dv$$

$$\sigma_x^2 = \text{var}[x] = \int (v-\mu_x)^2 f_x(v)\, dv = \iint (v-\mu_x)^2 f(v,w)\, dw\, dv$$

Covariance: $\sigma_{xy} = \text{cov}[x,y] = E[(x-\mu_x)(y-\mu_y)]$

$$= \iint (v-\mu_x)(w-\mu_y) f(v,w)\, dw\, dv$$

Corrrelation: $\rho_{xy} = \sigma_{xy}/(\sigma_x\sigma_y)$

Sum of two random variables

Let $z = x + y$

- Mean of z:

$$\begin{aligned}
E[z] = E[x+y] &= \iint (v+w) f(v,w)\, dw\, dv \\
&\stackrel{a}{=} \iint v\, f(v,w)\, dw\, dv + \iint w\, f(v,w)\, dw\, dv \\
&= E[x] + E[y] = \mu_x + \mu_y
\end{aligned}$$

Expectation of the sum is the sum of the expectations

- Variance of z

$$\begin{aligned}
\text{var}[z] = \text{var}[x+y] &= E\left[(x-\mu_x+y-\mu_y)^2\right] \\
&\stackrel{b}{=} E\left[(x-\mu_x)^2 + (y-\mu_y)^2 + 2(x-\mu_x)(y-\mu_y)\right] \\
&\stackrel{c}{=} E\left[(x-\mu_x)^2\right] + E\left[(y-\mu_y)^2\right] + E\left[2(x-\mu_x)(y-\mu_y)\right] \\
&= \text{var}[x] + \text{var}[y] + 2\,\text{cov}[x,y]
\end{aligned}$$

Linear function of two random variables

Test

Let x and y be random variables with means μ_x and μ_y, variances σ_x^2 and σ_y^2 and covariance σ_{xy}. Consider the linear transformation $z = a_1 x + a_2 y + b$ for constants a_1, a_2 and b. Calculate $E[z]$ and $\text{var}[z]$.

Answer

$$
\begin{aligned}
E[z] &= E[a_1 x + a_2 y + b] = E[a_1 x] + E[a_2 y] + E[b] \\
&= a_1 E[x] + a_2 E[y] + b = a_1 \mu_x + a_2 \mu_y + b \\
\text{var}[z] &= E\left[(z - E[z])^2\right] = E\left[\left(a_1(x-\mu_x) + a_2(y-\mu_y)\right)^2\right] \\
&= E\left[a_1^2 (x-\mu_x)^2\right] + E[(a_2^2 (y-\mu_y)^2] + E[2a_1 a_2 (x-\mu_x)(y-\mu_y)] \\
&= a_1^2 \sigma_x^2 + a_2^2 \sigma_y^2 + 2 a_1 a_2 \sigma_{xy}
\end{aligned}
$$

n random variables

Let y_i, $i = 1, 2, \ldots, n$, be random variables.

- Joint density $f(v)$, with v $(n \times 1)$

$$
P[a_1 \le y_1 \le b_1, \ldots, a_n \le y_n \le b_n] = \int_{a_1}^{b_1} \cdots \int_{a_n}^{b_n} f(v)\, dv_n \cdots dv_1
$$

- Marginal density for y_i by integral over all v_j, $j \neq i$

Expectation of n random variables

- n means, $\mu_i = E[y_i]$

$$
\mu = E[y] = \begin{pmatrix} E[y_i] \\ \vdots \\ E[y_n] \end{pmatrix} = \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_n \end{pmatrix}
$$

- n variances, $\sigma_i^2 = \text{var}[y_i]$ and
  $n(n-1)/2$ covariances, $\sigma_{ij} = \text{cov}[y_i, y_j] = \text{cov}[y_j, y_i] = \sigma_{ji}$

$$
\Sigma = E[(y-\mu)(y-\mu)'] = \begin{pmatrix} \sigma_1^2 & \sigma_{12} & \cdots & \sigma_{1n} \\ \sigma_{21} & \sigma_2^2 & \ddots & \sigma_{2n} \\ \vdots & \ddots & \ddots & \vdots \\ \sigma_{n1} & \sigma_{n2} & \cdots & \sigma_n^2 \end{pmatrix}
$$

Linear function of n random variables

$$
z = b + \sum_{i=1}^{n} a_i y_i = b + a'y
$$

$$
\text{Mean: } E[z] = E\left[b + \sum_{i=1}^{n} a_i y_i\right] = b + \sum_{i=1}^{n} a_i \mu_i = b + a'\mu
$$

$$
\begin{aligned}
\text{Variance: } \text{var}[z] &= E\left[(z - E[z])^2\right] = E\left[(a'(y-\mu))^2\right] \\
&= E\left[\left(\sum_{i=1}^{n} a_i (y_i - \mu_i)\right)^2\right] \overset{*}{=} E\left[\sum_{i=1}^{n}\sum_{j=1}^{n} a_i a_j (y_i - \mu_i)(y_j - \mu_j)\right] \\
&= \sum_{i=1}^{n}\sum_{j=1}^{n} a_i a_j E[(y_i - \mu_i)(y_j - \mu_j)] \\
&= \sum_{i=1}^{n}\sum_{j=1}^{n} a_i \sigma_{ij} a_j = a'\Sigma a \qquad (\sigma_{ii} = \sigma_i^2)
\end{aligned}
$$

Properties of Σ

Test

Let x be an n-variate random variable with $E[x] = \mu$ and $\text{var}[x] = \Sigma$. What properties does Σ have?

Answer

- Σ is symmetric, because $\sigma_{ij} = \text{cov}[y_i, y_j] = \text{cov}[y_j, y_i] = \sigma_{ji}$.
- Define $z = b + a'y$, then $\text{var}[z] = a'\Sigma a$. For all a, $\text{var}[z] \geq 0$. This means $a'\Sigma a \geq 0$, so Σ is positive semi-definite (PSD).

Linear transformations of random variables

Let z be a set of k random variables and y a set of n random variables, with

$$z = \underset{(k\times 1)}{b} + \underset{(k\times n)}{A}\, y$$

Mean vector: $\mu_z = E[z] = E[b + Ay] = b + A\mu_y$

Covariance matrix:

$$\begin{aligned}\Sigma_z =& E[(z-\mu_z)(z-\mu_z)'] \\ =& E\left[A(y-\mu_y)(A(y-\mu_y))'\right] \\ =& E\left[A(y-\mu_y)(y-\mu_y)'A'\right] \\ =& A\,E\left[(y-\mu_y)(y-\mu_y)'\right]A' \\ =& A\Sigma_y A'\end{aligned}$$

Training Exercise P.1

- Train yourself by making the training exercise (see the website).
- After making this exercise, check your answers by studying the webcast solution (also available on the website).