Multivariate distributions

Covariance

Definition **Covariance of X_i and X_j** is

$$\mathrm{Cov}(X_i, X_j) = \sigma_{X_i,X_j} = \mathrm{E}\big[(X_i - \mathrm{E}X_i)(X_j - \mathrm{E}X_j)\big] = \mathrm{E}X_iX_j - (\mathrm{E}X_i)(\mathrm{E}X_j)$$

Interpretation:

- $\mathrm{Cov}(X_i, X_j) > 0$ means that larger (or smaller) than average values of X_i tend to occur with larger (or smaller) than average values of X_j
- $\mathrm{Cov}(X_i, X_j) < 0$ means that larger (or smaller) than average values of X_i tend to occur with smaller (or larger) than average values of X_j

Relationships:

1. $\mathrm{Cov}(X_i, X_i) = \mathrm{Var}(X_i)$
2. $\mathrm{Cov}(X_i, X_j) = \mathrm{Cov}(X_j, X_i)$
3. $\mathrm{Cov}(aX_i + c, bX_j + d) = ab\mathrm{Cov}(X_i, X_j)$
4. $\mathrm{Var}(aX_i + bX_j + c) = a^2\mathrm{Var}(X_i) + b^2\mathrm{Var}(X_j) + 2ab\mathrm{Cov}(X_i, X_j)$

Proof

$$\begin{aligned}
& \mathrm{Var}(aX_i + bX_j + c) \\
= \; & \mathrm{E}\left(\left[aX_i + bX_j + c - \mathrm{E}(aX_i + bX_j + c)\right]^2\right) \\
= \; & \mathrm{E}\left(\left[aX_i + bX_j + c - a\mathrm{E}X_i - b\mathrm{E}X_j - c\right]^2\right) \\
= \; & \mathrm{E}\left(\left[a(X_i - \mathrm{E}X_i) + b(X_j - \mathrm{E}X_j)\right]^2\right) \\
= \; & \mathrm{E}\left(a^2(X_i - \mathrm{E}X_i)^2 + b^2(X_j - \mathrm{E}X_j)^2 + 2ab(X_i - \mathrm{E}X_i)(X_j - \mathrm{E}X_j)\right)
\end{aligned}$$

Multivariate distributions

Covariance (cont'd)

5. more generally,

$$\mathrm{Cov}\left(\underbrace{c+\sum_{i=1}^{n} a_i X_i}_{V}, \underbrace{d+\sum_{j=1}^{n} b_j X_j}_{W}\right) \overset{?}{=} \boxed{\sum_{i=1}^{n}\sum_{j=1}^{n} a_i b_j \mathrm{Cov}(X_i, X_j)} \quad \mathrm{I}$$

$$\overset{?}{=} \sum_{i=1}^{n} a_i b_i \mathrm{Var}(X_i) + 2 \sum_{1\le i<j\le n} a_i b_j \mathrm{Cov}(X_i, X_j) \quad \mathrm{II}$$

$$\boxed{V = c + \sum_{i=1}^{n} a_i X_i}$$

$$\mathbb{E}(V) = c + \sum_{i=1}^{n} a_i \mathbb{E}X_i$$

$$\mathbb{E}(W) = d + \sum_{j=1}^{n} b_j \mathbb{E}X_j$$

$$W = d + \sum_{j=1}^{n} b_j X_j$$

$$\mathrm{Cov}(V, W) \overset{\mathrm{def}}{=} \mathbb{E}\Big[\underbrace{(V-\mathbb{E}V)}_{*}\underbrace{(W-\mathbb{E}W)}_{**}\Big]$$

$$= \mathbb{E}\Big[\underbrace{\sum_{i=1}^{n} a_i (X_i - \mathbb{E}X_i)}_{*} \underbrace{\sum_{j=1}^{n} b_j (X_j - \mathbb{E}X_j)}_{**}\Big]$$

$$= \mathbb{E}\Big[\sum_{i=1}^{n}\sum_{j=1}^{n} a_i b_j (X_i - \mathbb{E}X_i)(X_j - \mathbb{E}X_j)\Big]$$

$$= \sum_{i=1}^{n}\sum_{j=1}^{n} a_i b_j \underbrace{\mathbb{E}\Big[(X_i - \mathbb{E}X_i)(X_j - \mathbb{E}X_j)\Big]}$$

$$= \sum_{i=1}^{n}\sum_{j=1}^{n} a_i b_j \mathrm{Cov}(X_i, X_j)$$

6. also, (4. above is a special case of the following)

$$\mathrm{Var}\left(c+\sum_{i=1}^{n} a_i X_i\right) = \sum_{i=1}^{n}\sum_{j=1}^{n} a_i a_j \mathrm{Cov}(X_i, X_j) \quad \mathrm{I}$$

$$= \sum_{i=1}^{n} a_i^2 \mathrm{Var}(X_i) + 2 \sum_{1\le i<j\le n} a_i a_j \mathrm{Cov}(X_i, X_j) \quad \mathrm{II}$$

$$\mathrm{Var}(\star) = \mathrm{Cov}(\star, \star)$$

$$\mathrm{Var}\left(c+\sum_{i=1}^{n} a_i X_i\right) = \mathrm{Cov}\left(c+\sum_{i=1}^{n} a_i X_i, c+\sum_{j=1}^{n} a_i X_j\right)$$

Multivariate distributions

Covariance: examples

Discrete example: $\mathbb{E}X = \mathbb{E}Y = 2, \quad \mathbb{E}XY = 50/12$

$$\mathrm{Cov}(X,Y) = \mathbb{E}XY - \mathbb{E}X\,\mathbb{E}Y = 50/12 - 4 = 1/6$$

Continuous example: $\circledast \; f(x,y) = \frac{1}{x}$ for $0 < y < x < 1$

$$\mathrm{Cov}(X,Y) = \mathbb{E}XY - \mathbb{E}X\,\mathbb{E}Y = 1/6 - 1/2 \cdot 1/4 = 1/24$$

$$\mathrm{Cov}(2X - 8Y, 4X + 4Y + \cancel{7}6)$$

$$= \mathrm{Cov}(2X - 8Y, 4X + 4Y)$$

$$= (2\times 4)\underbrace{\mathrm{Cov}(X,X)}_{1/12} + (2\times 4)\underbrace{\mathrm{Cov}(X,Y)}_{1/24} - (8\times 4)\underbrace{\mathrm{Cov}(Y,X)}_{1/24} - (8\times 4)\underbrace{\mathrm{Cov}(Y,Y)}_{7/144}$$

$$\mathrm{Cov}(X,X) = \mathrm{Var}(X) = \mathbb{E}X^2 - (\mathbb{E}X)^2 = \frac{1}{3} - \left(\frac{1}{2}\right)^2 = 1/12$$

$$\mathrm{Cov}(Y,Y) = \mathrm{Var}(Y) = \mathbb{E}Y^2 - (\mathbb{E}Y)^2 = 1/9 - (1/4)^2 = 7/144$$

Note: $\mathbb{E}(X+Y) = \iint (x+y) f_{X,Y}(x,y)\,dx\,dy$

$$= \iint x f_{X,Y}(x,y)\,dx\,dy + \iint y f_{X,Y}(x,y)\,dx\,dy$$

$$= \int x \left[\int f_{X,Y}(x,y)\,dy\right] dx + \int y \left[\int f_{X,Y}(x,y)\,dx\right] dy$$

$$= \underbrace{\int x f_X(x)\,dx + \int y f_Y(y)\,dy}_{= \mathbb{E}X + \mathbb{E}Y}$$

Multivariate distributions

Correlation

- Covariance has no natural scale

$\mathrm{Cov}(X,Y) = 42.6$

- **Correlation of X_i and X_j** is a standardized measure of association

$$\mathrm{Corr}(X_i, X_j) = \rho_{X_i,X_j} = \frac{\mathrm{Cov}(X_i, X_j)}{\sqrt{\mathrm{Var}(X_i)}\sqrt{\mathrm{Var}(X_j)}} = \frac{\mathrm{Cov}(X_i, X_j)}{\sigma_{X_i}\sigma_{X_j}}$$

$\mathrm{Var}\, X_i < \infty$, $\mathrm{Var}\, X_j < \infty$

Continuous Case: $f(x,y) = \frac{1}{x}$ if $0 < y < x < 1$

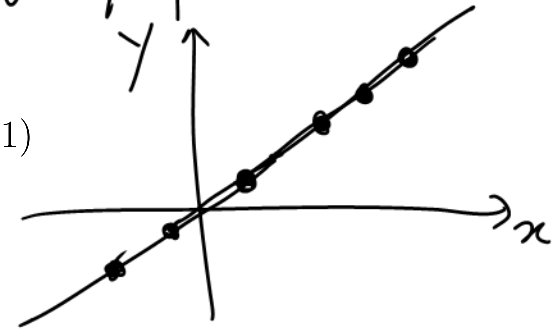
- Example

$$\mathrm{Corr}(X,Y) = \frac{\mathrm{Cov}(X,Y)}{\sqrt{\mathrm{Var}(X)}\sqrt{\mathrm{Var}(Y)}} = \frac{1/24}{\sqrt{1/12}\sqrt{7/144}} \approx .655$$

- **Theorem:**

$$\rho_{X,Y} = \left| \frac{\mathrm{Cov}(X,Y)}{\sqrt{\mathrm{Var}(X)}\sqrt{\mathrm{Var}(Y)}} \right| \leq 1$$

1. $-1 \leq \rho_{X,Y} \leq 1$
2. $\rho_{X,Y} = 1$ if and only if $P(X = aY + b) = 1$
   (where $a > 0$ if $\rho_{X,Y} = +1$ and $a < 0$ if $\rho_{X,Y} = -1$)



Multivariate distributions

Correlation (cont'd)

- *Proof*: $|\rho_{X,Y}| \le 1$ $\qquad E[(a+b)^2] = E[a^2+b^2+2ab]$
  - Define a function

$$h(t) = \mathrm{E}\Big(\underbrace{(X-\mathrm{E}X)t}_{a} + \underbrace{(Y-\mathrm{E}Y)}_{b}\Big)^2 = \underbrace{\sigma_X^2 t^2 + 2\mathrm{Cov}(X,Y)t + \sigma_Y^2}_{ax^2+bx+C} \quad \Delta = b^2-4ac$$

$$at^2 + bt + C$$

  - $h(t) \ge 0$, so either $h(\cdot)$ has one real root or two complex roots $\quad \frac{-b \pm \sqrt{\Delta}}{2a}$

roots of $h(t)$ are $\frac{-b\pm\sqrt{\Delta}}{2a}$.

$$\begin{cases} a = \sigma_X^2 \\ b = 2\,\mathrm{Cov}(X,Y) \\ c = \sigma_Y^2 \end{cases}$$

But $\Delta \le 0 \Rightarrow b^2 - 4ac \le 0$

$$\Rightarrow [2\mathrm{Cov}(X,Y)]^2 - 4(\sigma_X^2)(\sigma_Y^2) \le 0$$

$$\cancel{4}\mathrm{Cov}(X,Y)^2 \le \cancel{4}\sigma_X^2\sigma_Y^2$$

$$\frac{(\mathrm{Cov}(X,Y))^2}{\sigma_X^2\sigma_Y^2} = |\rho(X,Y)|^2 \le 1$$

$$|x|^2 \le a^2 \Leftrightarrow -a \le x \le a$$

$$\Leftrightarrow -1 \le \rho(X,Y) \le 1$$

Multivariate distributions

Moment generating functions

Definition: The **joint moment generating function** of $(X_1, \ldots, X_n)$ is

$$\longrightarrow \quad M_{X_1,\ldots,X_n}(t_1, \ldots, t_n) = \mathrm{E}e^{t_1X_1+\cdots+t_nX_n}, \quad t_1, \ldots, t_n \in \mathbb{R}$$

if the expectation exists for all $\underbrace{-h < t_1, \ldots, t_n < h}$ for some $h > 0$

- Joint mgf can provide univariate mgfs

$$M_{X_i}(t_i) = M_{X_1,\ldots,X_n}(t_1 = 0, \ldots, t_{i-1} = 0, t_i, t_{i+1} = 0, \ldots, t_n = 0)$$

$n=2 \quad X, Y$

$$M_{X,Y}(0, t_2) = \mathbb{E}\left[e^{0X + t_2 Y}\right] = \mathbb{E}\left[e^{t_2 Y}\right] = \boxed{M_Y(t_2)}$$

- Applications as before (more later):
  - characterizes distributions

    e.g., if $(X_1, \ldots, X_n)$ and $(Y_1, \ldots, Y_n)$ have the same mgfs, then these vectors have the same distribution

  - transformations

    e.g., mgf of $(a_1X_1, \ldots, a_nX_n)$ is $M_{X_1,\ldots,X_n}(a_1t_1, \ldots, a_nt_n)$

  - convergence (later)

  - moments, e.g,

$$\boxed{\mathrm{E}(X_i^q X_j^r X_k^s) = \frac{\partial^{q+r+s}}{\partial t_i^q \partial t_j^r \partial t_k^s} M_{X_1,\ldots,X_n}(t_1, \ldots, t_n)\Big|_{(t_1,\ldots,t_n)=(0,\ldots,0)}}$$

$$\mathbb{E}X^r = \frac{d^r}{dt^r} M_X(t)\Big|_{t=0}$$