Expected Value of a Function of Random Variables

There are k variables

Joint density

$$E[g(Y_1, Y_2, \ldots, Y_k)] =$$
$$\int_{y_k} \cdots \int_{y_2} \int_{y_1} \underbrace{g(y_1, y_2, \ldots, y_k)}_{\text{any ft of } Y_1, \ldots Y_k} \underbrace{f(y_1, y_2, \ldots, y_k)} dy_1 dy_2 \cdots dy_k.$$

a and b: constants

- $E(ag(Y_1, Y_2) + b) = aE[g(Y_1, Y_2)] + b$
- $E[\sum_{i=1}^{k} g_i(Y_1, Y_2)] = \sum_{i=1}^{k} E[g_i(Y_1, Y_2)]$
- $E[g(Y_1)h(Y_2)] = E[g(Y_1)]E[h(Y_2)]$ if Y_1 and Y_2 are independent. $\not\Longrightarrow$ indep.

$E(Y_1 Y_2) = E(Y_1) E(Y_2)$ if $Y_1 \perp Y_2$

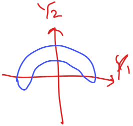
X, Y with joint pdf $f(x, y)$

$$E(X) = \iint x\, f(x,y)\, dx\, dy = \int x\, f(x)\, dx$$

marginal

Covariance of Two Random Variables

Y_1, Y_2: r.v.'s with means μ_1 and μ_2, respectively



Covariance only linear relationships measured.

· bet'n $-\infty$, ∞

· difficult to interpret the value itself.

$$Cov(Y_1, Y_2) = E[(Y_1 - \mu_1)(Y_2 - \mu_2)]. \quad = E(Y_1Y_2) - \mu_1\mu_2$$

· $Cov(Y, Y) = V(Y)$

Correlation coefficient

$$\rho = \frac{Cov(Y_1, Y_2)}{\sigma_1\sigma_2}. = \frac{Cov(Y_1, Y_2)}{\sqrt{V(Y_1)}\sqrt{V(Y_2)}}$$

$-1 \leq \rho \leq 1$

$\boxed{Y_1, Y_2\text{: independent, then } Cov(Y_1, Y_2) = 0.}$

- $V(Y_1 - aY_2) = Cov(Y_1 - aY_2,\ Y_1 - aY_2)$

$$= Cov(Y_1, Y_1) - 2a\,Cov(Y_1, Y_2) + a^2\,Cov(Y_2, Y_2)$$

$$= \sigma_1^2 - 2a\,Cov(Y_1, Y_2) + a^2\sigma_2^2 \geq 0 \quad \forall a$$

$$Cov(Y_1, Y_2)^2 - \sigma_1^2\sigma_2^2 \leq 0$$

$$\Rightarrow \quad -1 \leq \frac{Cov(Y_1, Y_2)}{\sigma_1\sigma_2} \leq 1.$$

$ba^2 + ca + d \geq 0$

$\forall a$

$c^2 - 4bd \leq 0$

- $Cov(aY_1 + b,\ cY_2 + d) = ac\,Cov(Y_1, Y_2)$
- $Cov(Y_1,\ Y_2 + Y_3) = Cov(Y_1, Y_2) + Cov(Y_1, Y_3)$

bilinear

- Why zero covariance does not imply independence.

indep
$$Y_1 = \begin{cases} 1 & \text{w.p. } \frac{1}{2} \\ -1 & \text{w.p. } \frac{1}{2} \end{cases} \qquad Y_2 = \begin{cases} 1 & \text{w.p. } \frac{1}{2} \\ -1 & \text{w.p. } \frac{2}{2} \end{cases}$$

$$\begin{cases} X_1 = Y_1 + Y_2 \\ X_2 = Y_1 - Y_2 \end{cases}$$

$$\boxed{X_1 = \begin{cases} -2 & \text{w.p. } \frac{1}{4} \\ 0 & \text{w.p. } \frac{1}{2} \\ 2 & \text{w.p. } \frac{1}{4} \end{cases}}$$

$$\boxed{X_2 = \begin{cases} -2 & \text{w.p. } \cdot \\ 0 & \cdot \\ 2 & \cdot \end{cases}}$$

Note

$Cov(X_1, X_2)$

$= Cov(Y_1 + Y_2,\ Y_1 - Y_2)$

$= V(Y_1) - V(Y_2)$

$= 0$

$P(X_1 = ?, X_2 = ??)$

$\neq P(X_1 = ?)P(X_2 =$

$\rightarrow$ You can show that X_1 and X_2 are dependent!!

Expected Value and Variance of Linear Functions of Random Variables

$E(Y_i)=\mu_i$ $\quad$ $E(X_j)=\xi_j$

$Y_1, Y_2, \ldots, Y_n$ and $X_1, X_2, \ldots, X_m$: r.v.'s with $E(Y_i) = \mu_i$ and $E(X_j) = \xi_j$.

$=$

$U_1 = \sum_{i=1}^{n} a_i Y_i$ and $U_2 = \sum_{j=1}^{m} b_j X_j$

a. $E(U_1) = \sum_{i=1}^{n} a_i \mu_i$

b. $V(U_1) = \sum_{i=1}^{n} a_i^2 V(Y_i) + 2 \sum \sum_{i<j} a_i a_j Cov(Y_i, Y_j)$

c. $Cov(U_1, U_2) = \sum_{i=1}^{n} \sum_{j=1}^{m} a_i b_j Cov(Y_i, X_j)$.

$V(Y_1+Y_2) = V(Y_1)+V(Y_2)+2Cov(Y_1,Y_2)$ vs. $V(Y_1)+V(Y_2)$

$V(Y_1-Y_2) = \quad " \quad\quad " \quad -2Cov(Y_1,Y_2)$

- $Y_1, \dots Y_n \overset{iid}{\sim}$ $E(Y_i) = \mu$, $V(Y_i) = \sigma^2$

- $E(\bar{Y}) = \mu$ by a.
- $V(\bar{Y}) = \frac{\sigma^2}{n}$ by b.

Now two important (Joint) multivariate distributions. .

Multinomial Probability Distribution ← Generalization of Binomial

Multinomial experiment

- n independent, identical trials.
- The outcome of each trial falls into one of k classes or cells. $k \geq 3$
- p_i: probability that the outcome of a single trial falls into cell $i = 1, 2, \ldots, k$, and $\boxed{\sum_{i=1}^{k} p_i = 1.}$
- Y_i: the number of trials for which the outcome falls into cell i. $\sum_{i=1}^{k} Y_i = n$. $(Y_1, \ldots, Y_k)$

Multinomial Probability Distribution

$Y_1, Y_2, \ldots, Y_k$: multinomial distribution PMF

$$p(y_1, y_2, \ldots, y_k) = \frac{n!}{y_1!y_2!\cdots y_k!} p_1^{y_1} p_2^{y_2} \cdots p_k^{y_k},$$

$\binom{n}{y_1 \cdots y_k}$

where $p_i > 0$ and $y_i = 0, 1, \ldots, n$ for $i = 1, 2, \ldots, k$, and $\sum_{i=1}^{k} p_i = 1$ and $\sum_{i=1}^{k} Y_i = n$.

- Its marginal distribution is binomial. $\rightarrow$ $Y_1 \sim B(n, p_1)$ $\vdots$ $Y_k \sim B(n, p_k)$

  $Y_1 + Y_2 \sim B(n, p_1 + p_2)$

- $E(Y_i) = np_i$, $V(Y_i) = np_iq_i$.
- $Cov(Y_s, Y_t) = \boxed{-np_sp_t}$, if $t \neq s$.

$V(Y_1 + Y_2) = V(Y_1) + V(Y_2) + 2\,Cov(Y_1, Y_2)$

$n(p_1 + p_2)(1 - p_1 - p_2) = np_1q_1 + np_2q_2 + 2\,Cov(Y_1, Y_2)$

$\star \Rightarrow \boxed{Cov(Y_1, Y_2) = -np_1p_2}$

Ex 123

Fires

73%	Home	$\to$ \$20,000	
20%	Apt	$\to$ \$10,000	cost
7%	other	$\to$ \$2000	

$n = 4$ fires

$$E(\text{Total Cost}) = E(20000Y_1 + \cdots + 2000Y_3) = \cdots\cdots$$

$$V(\text{Total Cost}) = V(20000Y_1 + \cdots + 2000Y_3) = 20000^2 \cdot 4 \cdot (.73)(1-.73)$$

$$+ \cdots\cdots$$

$$+ 2(20{,}000)(10000)(-4\cdot(.73)(.20))$$

$$\vdots$$

$$= \cdots\cdots$$

$(Y_1, Y_2, Y_3) \sim \text{Multinomial}(4, .73, .20, .7)$

Y_1: # Home fires out of 4

Bivariate Normal Distribution

"Y_1 & Y_2 are jointly normal"

$\rightarrow N_2(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$

$$N_2\left(\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} = \underline{\mu},\ \underbrace{\begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}}_{= \Sigma_{2\times 2}} \right)$$

Y_1 and Y_2 are bivariate normal if

$$f(y_1, y_2) = \frac{e^{-Q/2}}{2\pi\underbrace{\sigma_1\sigma_2\sqrt{1-\rho^2}}_{= |\Sigma|}}, \quad -\infty < y_1, y_2 < \infty$$

where

$$Q = \frac{1}{1-\rho^2}\left[\frac{(y_1-\mu_1)^2}{\sigma_1^2} - 2\rho\frac{(y_1-\mu_1)(y_2-\mu_2)}{\sigma_1\sigma_2} + \frac{(y_2-\mu_2)^2}{\sigma_2^2}\right].$$

$$f(y_1, y_2) = \frac{1}{2\pi|\Sigma|^{\frac{1}{2}}} \exp\left[-\frac{1}{2}(\underline{y}-\underline{\mu})'\Sigma^{-1}(\underline{y}-\underline{\mu}) \right]$$

- $\text{Corr}(Y_1, Y_2) = \rho = 0 \implies Y_1 \perp Y_2$

Why?

DIY $f(y_1, y_2) = \frac{1}{\sqrt{2\pi}\sigma_1} \exp\left(-\frac{(y_1-\mu_1)^2}{2\sigma_1^2}\right) \times \frac{1}{\sqrt{2\pi}\sigma_2} \exp\left(-\frac{(y_2-\mu_2)^2}{2\sigma_2^2}\right)$

$= f(y_1) \times f(y_2)$

- Marginal

$Y_1 \sim N(\mu_1, \sigma_1^2)$

$Y_2 \sim N(\mu_2, \sigma_2^2)$

- Conditional

$Y_1 \mid Y_2 = y_2 \sim$