Covariances

Thursday, 24 February 2022 11:49

X & Y

p.d.f: $f(X,Y) \rightarrow p(x) \leftarrow$ Marginal dist w.r.t 'X'

$$p(X) = \int_X f(X,Y)\,dY$$

$$p(Y) = \int_Y f(X,Y)\,dX \leftarrow$$ Marginal dist. with respect to Y.

Ex: Roll two Dice. let X be the value on the first dice and T be the total on both sides.

Joint probability table.

$$X = \{1,2,3,4,5,6\}$$

$$T = \{2,3,4,\dots,12\}$$

X \ T	2	3	4	5	6	7	8	9	10	11	12	$P(x_i)$
1	1/36	1/36	1/36	1/36	1/36	1/36	0	0	0	0	0	1/6
2	0	1/36	1/36	1/36	1/36	1/36	1/36	0	0	0	0	1/6
3	0	0	1/36	1/36	1/36	1/36	1/36	1/36	0	0	0	1/6
4	0	0	0	1/36	1/36	1/36	1/36	1/36	1/36	0	0	1/6
5	0	0	0	0	1/36	1/36	1/36	1/36	1/36	1/36	0	1/6
6	0	0	0	0	0	1/36	1/36	1/36	1/36	1/36	1/36	1/6
$P(t_j)$	1/36	2/36	3/36	4/36	5/36	6/36	5/36	4/36	3/36	2/36	1/36	

Event: '$X=3$' $\rightarrow$ third row.

$$\underbrace{P(X=3) = 1/6 \qquad P(T=5) = 4/36}_{\text{Marginal probabilities.}}$$

Continuous Case:

(i) X has range $[0, 1/2]$, $Y \in [0,1]$

$f(x,y) = 96\,x^2 y^3$ then X + Y ...

independent. → Covariance!

Marginal densities : $f_X(x) = 24x^2$

Marginal density $f_Y(y) = 4y^3$.

$f(x,y) = f_X(x)\,f_Y(y) = (24x^2)(4y^3) = 96x^2y^3$

(2) $f(x,y) = 1.5\,(x^2+y^2)$ over the unit square.

$X \in [0,1] \quad Y \in [0,1]$.

X & Y are independent?

No : $f(x,y) \neq f_X(x)\,f_Y(y)$.

$f(x,y) = 96\,x^2y^3. \qquad X \in [0, 1/2]$

$\qquad\qquad\qquad\qquad\qquad Y \in [0, 1]$.

$$f_X(x) = \int_0^1 dy\, f(x,y)$$

$$= \int_0^1 dy \cdot 96x^2y^3 = \frac{96x^2y^4}{4}\Big|_0^1 = 24x^2.$$

$$f_Y(y) = \int_0^{1/2} dx \cdot 96x^2y^3 = \frac{96x^3}{3}y^3\Big|_0^{1/2}$$

$$= \frac{96}{3}\left(\frac{1}{2}\right)^3 y^3 = \frac{96}{24}y^3 = 4y^3.$$

$$\boxed{f(x,y) = (4y^3)(24x^2) = 96x^2y^3}$$

Covariance:

Covariance:

Covariance is a measure of how much two random variables vary together.

Def: Suppose X & Y are random variables with mean μ_X and μ_Y. The covariance of X and Y is defined by

$$Cov(X,Y) = E\big((X-\mu_X)(Y-\mu_Y)\big)$$

Properties of Covariance:

1. $Cov(aX+b,\ cY+d) = ac\,Cov(X,Y)$ for constant a,b,c,d.
2. $Cov(X_1+X_2, Y) = Cov(X_1,Y) + Cov(X_2,Y)$.
3. $Cov(X,X) = Var(X)$.
4. $Cov(X,Y) = E(XY) - \mu_X \mu_Y$.
5. $\boxed{Var(X+Y) = Var(X) + Var(Y) + 2\,Cov(X,Y)}$.

$Var(X+Y) = Var(X) + Var(Y)$ if and only if X & Y are independent.

when X & Y are independent: $Cov(X,Y)=0$.

6. If X & Y are independent Then $Cov(X,Y)=0$.

WARNING: The Converse is false.

Zero covariance does not always imply independence.

If $X=Y$:

$cov(X,Y) = E(X \cdot X) - \mu_X \mu_X$.

$Var(X) = E(X^2) - \mu_X^2$. (property 4)

$X \rightarrow aX+b = X'$ a, b are constant.

$Var(aX+b) = a^2 Var(X)$.

$= Y' = cY+d$.

$$Var(X'+Y') = Var(X') + Var(Y') + 2cov(X',Y')$$

$$= Var(aX+b) + Var(cY+d) + 2cov(aX+b, cY+d).$$

$$= a^2 Var(X) + c^2 Var(Y) + 2ac\, Cov(X,Y).$$

Sums and Integrals for Computing Covariance:

Discrete Case:

X and Y have joint pdf $p(x_i, y_j)$

$$Cov(X,Y) = \sum_{i=1}^{n} \sum_{j=1}^{m} p(x_i\, y_j)(x_i - \mu_X)(y_j - \mu_Y)$$

$$= \underbrace{\left(\sum_{i=1}^{n} \sum_{j=1}^{m} p(x_i\, y_j)\, x_i\, y_j \right)}_{E(XY)} - \mu_X \mu_Y.$$

Step: $p(x_i\, y_j) x_i\, y_j - p(x_i\, y_j) x_i \mu_Y - p(x_i\, y_j) y_j \mu_X + \mu_X \mu_Y\, p(x_i\, y_j)$

Continuous Case: $X, Y \rightarrow$ joint pdf $f(x,y)$

$$Cov(X,Y) = \int^{d} \int^{b} (x - \mu_X)(y - \mu_Y) f(x,y)\, dx\, dy.$$

$$= \left(\underbrace{\int_a^b \int_c^d f(x,y)\, xy \, dx\, dy}_{E(XY)} \right) - \mu_X \mu_Y$$

Example:

Flip a fair coin 3 times. let X be the number of heads in the first 2 flips. and let Y be the number of heads in the last 2 flips. Compute $Cov(X,Y)$.

$X \to \{0, 1, 2\}$. $\quad X=0 \leftarrow$ T T T

$Y \in \{0, 1, 2\}$. $\quad X=1 \begin{cases} \text{H T T} \\ \text{T H T} \end{cases}$

When you flip coins 3 times 8 possibilities: $\{HHH, HHT, \dots\}$. $\quad X=2:$ H H T, H H H

$X \backslash Y$	0	1	2	$P(x_i)$
0	1/8	1/8	0	1/4
1	1/8	2/8	1/8	1/2
2	0	1/8	1/8	1/4
$P(y_j)$	1/4	1/2	1/4	1

$X=0$, $Y=0$: T T T

$X=0$, $Y=1$: T T H

$X=0$, $Y=2$: T T _ X

$X=1$, $Y=1$: H T H, T H T

$E(X) = 1$

$E(Y) = 1$

$$Cov(X,Y) = E\left((X-\mu_X)(Y-\mu_Y)\right)$$
$$= \sum_{i,j} p(x_i, y_j)(x_i - 1)(y_j - 1)$$

$x_i = 1$ or $y_j = 1$ or $p = 0$.

$$Cov(X,Y) = \frac{1}{8}(0-1)(0-1) + \frac{1}{8}(2-1)(2-1) = \frac{1}{4}.$$

$$E(XY) = 1 \cdot \frac{2}{8} + 2 \cdot \frac{1}{8} + 4\frac{1}{8} = \frac{5}{4}$$

So

$$Cov(X,Y) = E(XY) - \mu_X \mu_Y.$$
$$= \frac{5}{4} - (1)(1) = \frac{5}{4} - 1 = \frac{1}{4}.$$

Next lecture:

- Covariance and correlations
- Bivariat Gaussian distribution.