Expectation involving two r.v.s.

Covariance, & Correlation Coefficient.

Conditional expectation, Law of total expectation

Conditional variance, Law of total variance.

Expectation involving two r.v.s.

- Let $(X,Y) \sim f_{X,Y}(x,y)$ and let $g(x,y)$ be a function of x and y.

  The <u>expectation of $g(X,Y)$</u> is given by

$$\mathbb{E}[g(X,Y)] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y)\, f_{X,Y}(x,y)\, dx\, dy.$$

  eg: $g(X,Y)$ may be X, Y, X^2, $X+Y$. etc.

- The <u>covariance</u> of X and Y is defined as

$$\begin{aligned}\mathrm{Cov}(X,Y) &= \mathbb{E}[(X-\mathbb{E}X)(Y-\mathbb{E}Y)] \\ &= \mathbb{E}[XY] - \mathbb{E}X\,\mathbb{E}Y.\end{aligned}$$

- X and Y are said to be <u>uncorrelated</u> if $\mathrm{Cov}(X,Y)=0$.

- The <u>correlation coefficient</u> of X and Y is defined as $\rho_{X,Y} = \dfrac{\mathrm{Cov}(X,Y)}{\sqrt{\mathrm{Var}(X)}\cdot\sqrt{\mathrm{Var}(Y)}}$.

Example : Mean, variance, covariance of a pair of r.v.s

Find $\mathbb{E}X$, $Var(X)$, and $Cov(X, Y)$ for $(X, Y) \sim f(x,y)$ where

$$f(x,y) = \begin{cases} 2, & x \geq 0,\ y \geq 0,\ x+y \leq 1. \\ 0. & \text{o.w.} \end{cases}$$

Solution

$$\mathbb{E}X = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x f(x,y)\,dx\,dy = \int_0^1\int_0^{1-x} 2x\,dy\,dx$$

$$= 2\int_0^1 (1-x)x\,dx = 2\left(\frac{1}{2}-\frac{1}{3}\right) = \frac{1}{3}$$

To find $Var(X) = \mathbb{E}(X^2) - (\mathbb{E}X)^2$, we first find $\mathbb{E}(X^2)$.

$$\mathbb{E}(X^2) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x^2 f(x,y)\,dx\,dy = \int_0^1\int_0^{1-x} 2x^2\,dy\,dx$$

$$= 2\int_0^1 (1-x)x^2\,dx = 2\left(\frac{1}{3}-\frac{1}{4}\right) = \frac{1}{6}$$

$$\Rightarrow Var(X) = \mathbb{E}[X^2) - (\mathbb{E}X)^2 = \frac{1}{6} - \frac{1}{9} = \frac{1}{18}$$

To find $\mathrm{Cov}(XY) = \mathbb{E}[XY] - \mathbb{E}X \cdot \mathbb{E}Y$, we find $\mathbb{E}Y$ and $\mathbb{E}[XY]$.

$$\mathbb{E}[Y] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} y f(x,y)\,dx\,dy = \int_0^1\int_0^{1-y} 2y\,dx\,dy$$

$$= 2\int_0^1 y(1-y)\,dy = 2\left(\frac{1}{2}-\frac{1}{3}\right) = \frac{1}{3}$$ (Or by symmetry, $\mathbb{E}Y = \mathbb{E}X$.)

$$\mathbb{E}[XY] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy\,f(x,y)\,dx\,dy$$

$$= \int_0^1\int_0^{1-y} 2xy\,dx\,dy$$

$$= \int_0^1 y(1-y)^2\,dy$$

$$= \frac{1}{12}$$

$$\Rightarrow \mathrm{Cov}(X,Y) = \mathbb{E}[XY] - \mathbb{E}X\cdot\mathbb{E}Y = \frac{1}{12} - \frac{1}{3}\cdot\frac{1}{3} = -\frac{1}{36}.$$

Uncorrelation v.s. Independence

- If X and Y are independent, then they are uncorrelated.

<u>proof</u>: $\mathbb{E}[XY] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy\, f(x,y)\,dx\,dy$

$$\overset{\text{independence}}{=} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy\, f(x) f(y)\, dx\, dy.$$

$$= \int_{-\infty}^{\infty} y f(y) \left[\int_{-\infty}^{\infty} x f(x)\,dx\right] dy$$

$$= \int_{-\infty}^{\infty} y f(y)\, \mathbb{E}X\, dy$$

$$= \mathbb{E}X \cdot \mathbb{E}Y.$$

$\Rightarrow \mathrm{Cov}(X,Y) = \mathbb{E}(X\cdot Y) - \mathbb{E}X\cdot\mathbb{E}Y = 0 \Rightarrow X, Y$ are uncorrelated.

- X and Y uncorrelated does NOT necessarily imply they are independent!

Example : Uncorrelation $\not\Rightarrow$ Independence.

Let $X, Y \in \{-2, -1, 1, 2\}$ such that

$$p(x,y) = \begin{cases} 2/5, & (x,y) = (1,1) \text{ or } (-1,-1) \\ 1/10, & (x,y) = (2,-2) \text{ or } (-2,2) \\ 0, & \text{o.w.} \end{cases}$$

Are X and Y independent? Are they uncorrelated?

Solution.

X and Y are NOT independent. eg: $P_{X,Y}(1,-1) = 0 \neq P_X(1) \cdot P_Y(-1) = \frac{4}{25}$

$$\mathbb{E}X = \sum_x x P_X(x) = P_X(1) - P_X(-1) + 2P_X(2) - 2P_X(-2) = \frac{2}{5} - \frac{2}{5} + \frac{2}{10} - \frac{2}{10} = 0$$

$\mathbb{E}Y = 0$ (by symmetry)

$$\mathbb{E}[XY] = \sum_x \sum_y xy P_{XY}(x,y) = 1^2 \cdot \frac{2}{5} + (-1)^2 \frac{2}{5} + 2(-2)\frac{1}{10} + (-2)2 \cdot \frac{1}{10} = 0$$

$\Rightarrow \operatorname{Cov}(X,Y) = \mathbb{E}(XY) - \mathbb{E}X \cdot \mathbb{E}Y = 0$, X and Y are uncorrelated.

Conditional probability given an event.

- <u>Conditioning on an event</u>: Let X be a r.v. with pmf $P_X(x)$.

Let A be an event s.t. $\mathbb{P}(X \in A) \neq 0$ We define the conditional pmf of X given $X \in A$ as

$$P_{X|A}(x) = \mathbb{P}(X = x \mid X \in A) = \frac{\mathbb{P}(X = x, X \in A)}{\mathbb{P}(X \in A)} = \begin{cases} \frac{P_X(x)}{\mathbb{P}(X \in A)} & \text{if } x \in A. \\ 0. & \text{o.w.} \end{cases}$$

- Note that $P_{X|A}(x)$ is a pmf on X.

- Similary, if X is a continuous r.v. with pdf $f_X(x)$, then

$$f_{X|A}(x) = \begin{cases} \frac{f_X(x)}{\mathbb{P}(X \in A)}, & \text{if } x \in A \\ 0, & \text{o.w.} \end{cases}$$

- Similarly, $f_{X|A}(x)$ is a pdf on X.

Conditional expectation given an event

- We define <u>conditional expectation $g(X)$ given $X \in A$</u> as

$$\mathbb{E}[g(X) \mid A] = \int_{-\infty}^{\infty} g(x) f_{X|A}(x)\,dx$$

recall $\mathbb{E}[g(X)] = \int_{-\infty}^{\infty} g(x) f_X(x)\,dx$

- eg 1: $g(x) = x$. $\mathbb{E}[X \mid A] = \int_{-\infty}^{\infty} x f_{X|A}(x)\,dx$.

  $\mathbb{E}[X] = \int_{-\infty}^{\infty} x f_X(x)\,dx$

  eg 2. $g(x) = x^2$ $\mathbb{E}[X^2 \mid A] = \int_{-\infty}^{\infty} x^2 f_{X|A}(x)\,dx$.

  $\mathbb{E}[X^2] = \int_{-\infty}^{\infty} x^2 f_X(x)\,dx$.

- <u>Law of total expectation</u>: Let $X \sim f_X(x)$. and $A_1, A_2, \ldots, A_n$ be disjoint nonzero probability events with $\mathbb{P}(\bigcup_{i=1}^{n} A_i) = \sum_{i=1}^{n} \mathbb{P}(A_i) = 1$. Then

$$\mathbb{E}[g(X)] = \sum_{i=1}^{n} \mathbb{P}(X \in A_i)\, \mathbb{E}[g(X) \mid A_i].$$

Law of total expectation. (Proof)

Pf: First note that by the law of total probability

$$f_X(x) = \sum_{i=1}^{n} \mathbb{P}(A_i) f_{X|A_i}(x)$$

Therefore

$$\mathbb{E}[g(X)] = \int_{-\infty}^{\infty} g(x) f_X(x)\, dx$$

$$= \int_{-\infty}^{\infty} g(x) \sum_{i=1}^{n} \mathbb{P}(A_i) f_{X|A}(x)\, dx.$$

$$= \sum_{i=1}^{n} \mathbb{P}(A_i) \int_{-\infty}^{\infty} g(x) f_{X|A_i}(x)\, dx$$

$$= \sum_{i=1}^{n} \mathbb{P}(A_i)\, \mathbb{E}[g(X) \mid A_i] \qquad \square$$

- This result is useful in computing expectation by devide-and-conquer.

Conditioning on a r.v.

- Let $(X,Y) \sim f_{X,Y}(x,y)$, and $f_Y(y) \neq 0$. We define the <u>conditional expectation of any function $g(X,Y)$ w.r.t. $f_{X|Y}(x|y)$</u> as

$$\mathbb{E}[g(X,Y) \mid Y=y] = \int_{-\infty}^{\infty} g(x,y) f_{X|Y}(x|y)\, dx .$$

- Ex 1: If $g(X,Y) = X$, then the conditional expectation of X given $Y=y$ is

$$\mathbb{E}[X \mid Y=y] = \int_{-\infty}^{\infty} x f_{X|Y}(x|y)\, dx$$

recall $\mathbb{E}X = \int_{-\infty}^{\infty} x f_X(x)\, dx$.

Ex 2: If $g(X,Y) = XY$, then

$$\mathbb{E}[XY \mid Y=y] = \int_{-\infty}^{\infty} xy f_{X|Y}(x|y)\, dx = y \int_{-\infty}^{\infty} x f_{X|Y}(x|y)\, dx$$

$$= y\, \mathbb{E}[X \mid Y=y].$$

Example: Conditional expectation

Let $f_{X,Y}(x,y) = \begin{cases} 2 & \text{if } x \geq 0,\ y \geq 0,\ x+y \leq 1 \\ 0 & \text{o.w.} \end{cases}$

Find $\mathbb{E}[X \mid Y=y]$ and $\mathbb{E}[XY \mid Y=y]$

Solution.

We already know that $f_{X|Y}(x|y) = \begin{cases} \frac{1}{1-y} & \text{if } x \geq 0,\ y \geq 0,\ x+y \leq 1. \\ 0, & \text{o.w.} \end{cases}$

Thus $\mathbb{E}[X \mid Y=y] = \int_0^{1-y} x\, f_{X|Y}(x|y)\, dx = \int_0^{1-y} \frac{x}{1-y}\, dx = \frac{1-y}{2}, \quad 0 \leq y \leq 1.$

$\mathbb{E}[XY \mid Y=y] = y\, \mathbb{E}[X \mid Y=y] = \frac{y(1-y)}{2}, \quad 0 \leq y \leq 1.$

- Rmk: Note that both $\mathbb{E}[X \mid Y=y]$ and $\mathbb{E}[XY \mid Y=y]$ are functions in y.

Conditional Expectation as a r.v.

- We define the <u>conditional expectation of $g(X,Y)$ given Y</u>, written as $\mathbb{E}[g(X,Y)\mid Y]$, as the random variable that takes value $\mathbb{E}[g(X,Y)\mid Y=y]$ when $Y=y$.

- $\boxed{\mathbb{E}[g(X,Y)\mid Y] \text{ is a function of the r.v. } Y.}$

eg: From the previous example,

$$\mathbb{E}[X\mid Y] = \frac{1-Y}{2}. \qquad \text{and} \qquad \mathbb{E}[XY\mid Y] = \frac{Y(1-Y)}{2}$$

When $Y=y$, $\mathbb{E}[X\mid Y=y] = \frac{1-y}{2}$. $\mathbb{E}[XY\mid Y=y] = \frac{y(1-y)}{2}$.

Law of total expectation (Iterated expectation)

- For any function $g(X, Y)$,

$$\boxed{\mathbb{E}[g(X,Y)] = \mathbb{E}_Y\left[\mathbb{E}_{X|Y}[g(X,Y)\,|\,Y]\right]}$$

where the inner expectation $\mathbb{E}_{X|Y}$ is w.r.t. $f_{X|Y}(x|y)$ and the outer expectation $\mathbb{E}_Y$ is w.r.t. $f_Y(y)$.

- Pf:

$$\begin{aligned}
\mathbb{E}_Y\left[\mathbb{E}_{X|Y}[g(X,Y)\,|\,Y]\right] &= \int_{-\infty}^{\infty} \mathbb{E}_{X|Y}[g(X,Y)\,|\,Y=y]\, f_Y(y)\, dy \\
&= \int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} g(x,y)\, f_{X|Y}(x|y)\, dx \right) f_Y(y)\, dy \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x,y)\, f_{XY}(x,y)\, dx\, dy \\
&= \mathbb{E}[g(X,Y)]. \qquad \square
\end{aligned}$$

Example: Infer the unknown bias in choosing lab problems.

- Θ is drawn uniform at random in $[0,1]$.
- Given $\Theta = \theta$, each problem is chosen with prob. θ, indep. of each other.
- n problems in total (fixed). K problems were chosen
- Find $\mathbb{E}[K]$

Solution

We already know that $K \mid \Theta = \theta \sim \text{Binom}(n, \theta)$. $\Rightarrow \mathbb{E}[K \mid \Theta = \theta] = n\theta$

By the law of total expectation,

$$\mathbb{E}[K] = \mathbb{E}_{\Theta}\big[\mathbb{E}[K \mid \Theta]\big] = \mathbb{E}[n\Theta] = \frac{n}{2}.$$

Rmk: Law of total expectation can be useful when the conditional expectation is easier to compute than the unconditional expectation.

Conditional Variance.

- Let X and Y be two r.v.s We define the <u>conditional variance of X given $Y=y$</u> as the variance of X w.r.t. $f_{X|Y}(x|y)$, i.e.,

$$\mathrm{Var}[X|Y=y] = \mathbb{E}[X^2|Y=y] - (\mathbb{E}[X|Y=y])^2.$$

- Define $\mathrm{Var}[X|Y]$ as the r.v. that takes value $\mathrm{Var}[X|Y=y]$ when $Y=y$.
- <u>$\mathrm{Var}[X|Y]$ is a function of the r.v. Y.</u>

Law of total variance

- Let X, Y be two r.v.s. We have

$$\boxed{\mathrm{Var}[X] = \mathbb{E}[\mathrm{Var}[X|Y]] + \mathrm{Var}[\mathbb{E}[X|Y]].}$$

- Pf: The expectation of the r.v. $\mathrm{Var}[X|Y]$ is

$$\mathbb{E}_Y[\mathrm{Var}[X|Y]] = \mathbb{E}_Y[\mathbb{E}[X^2|Y]] - \mathbb{E}_Y[(\mathbb{E}[X|Y])^2]$$

law of total expectation

$$= \mathbb{E}[X^2] - \mathbb{E}_Y[(\mathbb{E}[X|Y])^2].$$

The variance of the r.v. $\mathbb{E}[X|Y]$ is

$$\mathrm{Var}[\mathbb{E}[X|Y]] = \mathbb{E}_Y[(\mathbb{E}[X|Y])^2] - (\mathbb{E}_Y[\mathbb{E}[X|Y]])^2$$

law of total expectation

$$= \mathbb{E}_Y[(\mathbb{E}[X|Y])^2] - (\mathbb{E}X)^2$$

□.

Example: Infer the unknown bias in choosing lab problems.

- Θ is drawn uniform at random in $[0,1]$.
- Given $\Theta=\theta$, each problem is chosen with prob. θ, indep. of each other.
- n problems in total (fixed). K problems were chosen
- Find $\mathrm{Var}[K]$

$$\boxed{\text{For } X \sim \mathrm{Unif}[a,b],\ \mathbb{E}X = \frac{a+b}{2},\ \mathrm{Var}X = \frac{(b-a)^2}{12}}$$

Solution

$$K \mid \Theta=\theta \sim \mathrm{Binom}(n,\theta) \Rightarrow \mathbb{E}[K \mid \Theta=\theta] = n\theta \text{ and } \mathrm{Var}[K \mid \Theta=\theta] = n\theta(1-\theta)$$

By the law of total variance,

$$\mathrm{Var}[K] = \mathbb{E}[\mathrm{Var}[K \mid \Theta]] + \mathrm{Var}[\mathbb{E}[K \mid \Theta]] = \mathbb{E}[n\Theta(1-\Theta)] + \mathrm{Var}[n\Theta]$$

$$= n\,\mathbb{E}[\Theta] - n\,\mathbb{E}[\Theta^2] + n^2\,\mathrm{Var}[\Theta].$$

$$= n\,\mathbb{E}[\Theta] - n\left[\mathrm{Var}[\Theta] + (\mathbb{E}[\Theta])^2\right] + n^2\,\mathrm{Var}[\Theta]$$

$$= \frac{n}{6} + \frac{n^2}{12}$$