Lecture 16

Expectation of a function

Some rules

Functions of many variables

Multiplication rule

Expectation as prediction

Expectation of a function

$$E\{g(X)\} = \sum_x g(x)P\{X = x\}$$

- Here g is <u>any</u> function satisfying $\sum_x |g(x)|P\{X = x\} < \infty$.
- Recall: $E\{X\} = \sum_x xP\{X = x\}$.
- Therefore, in general, $E\{g(X)\} \neq g(E\{X\})$.
- If $g(x) = x^k$, $k = 1, 2, 3, \ldots$ then we get the <u>k-th moment of X</u>:

$$\boxed{m_k \equiv E\left\{X^k\right\}.}$$

- The definition makes sense!

The definition makes sense!

$$E\{g(X)\} \;=\; \sum_x g(x)P\{X = x\}$$

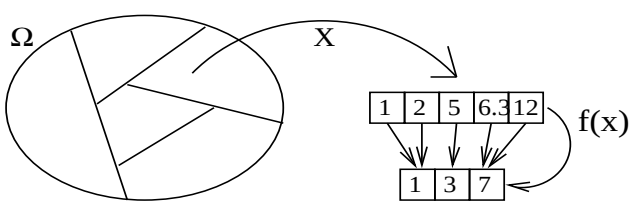
- $Y = g(X)$ is a random variable, and

$$p_Y(y) \;=\; P\{Y = y\} \;=\; \sum_{x:g(x)=y} P\{X = x\}.$$

- Therefore, $E\{Y\} = \sum_y y\, p_Y(y)$

$$E\{Y\} \;=\; \sum_y y\, p_Y(y) \;=\; \sum_y y \sum_{x:g(x)=y} P\{X = x\}$$

$$= \;\sum_x g(x)P\{X = x\} \;=\; E\{g(X)\}$$



X uniform on $[1, 2, \ldots, n]$.

$$E\{X\} = \sum_{x=1}^{n} \frac{1}{n} x = \frac{1}{n} \frac{n(n+1)}{2} = \frac{n+1}{2} \qquad 3.5$$

$$E\{X^2\} = \sum_{x=1}^{n} \frac{1}{n} x^2 = \frac{1}{n} \frac{n(n+1)(2n+1)}{6} = \frac{(n+1)(2n+1)}{6} \qquad \frac{91}{6}$$

$X \sim \mathrm{Bin}(n, p)$.

$$E\{X\} = \sum_{x=0}^{n} x \binom{n}{x} p^x (1-p)^{n-x} = np$$

$$E\{X^2\} = \sum_{x=0}^{n} x^2 \binom{n}{x} p^x (1-p)^{n-x} = n^2p^2 + npq$$

$$E\left\{X + X^2 + \sqrt{X}\right\} = \sum_{x=0}^{n} \left[x + x^2 + \sqrt{x}\right] \binom{n}{x} p^x (1-p)^{n-x} = ??$$

Some rules

$$E\{g(X)\} \;=\; \sum_x g(x)\; P\{X = x\}$$

- Constant factors: $E\{cX\} \;=\; cE\{X\}$.
- Constant variables: $E\{c\} \;=\; c$.
- Linear functions: (where $E\{g(X)\}=g(E\{X\})$!)

$$E\{aX + b\} \;=\; aE\{X\} + b.$$

- Addition rule: (used, but never proven)

$$E\{X_1 + X_2 + \ldots + X_n\} \;=\; E\{X_1\} + E\{X_2\} + \ldots + E\{X_n\}$$

Functions of many variables:

$$E\{g(X,Y)\} = \sum_x \sum_y g(x,y) P\{X = x, Y = y\}$$

PROOF: Set $Z = g(X,Y)$.

$$\begin{aligned} E\{g(X,Y)\} = E\{Z\} &= \sum_z z \; P\{Z = z\} \\ &= \sum_z z \left[\sum_{(x,y):g(x,y)=z} P\{X = x, Y = y\} \right] \\ &= \sum_{(x,y)} g(x,y) P\{X = x, Y = y\} \end{aligned}$$

Example: Tossing two fair dice

$$\begin{bmatrix} (1,1) & (1,2) & (1,3) & (1,4) & (1,5) & (1,6) \\ (2,1) & (2,2) & (2,3) & (2,4) & (2,5) & (2,6) \\ (3,1) & (3,2) & (3,3) & (3,4) & (3,5) & (3,6) \\ (4,1) & (4,2) & (4,3) & (4,4) & (4,5) & (4,6) \\ (5,1) & (5,2) & (5,3) & (5,4) & (5,5) & (5,6) \\ (6,1) & (6,2) & (6,3) & (6,4) & (6,5) & (6,6) \end{bmatrix}$$

$$\begin{aligned} E\{XY^2\} &= \frac{1}{36} \times \left[1 \times 1^2 + 1 \times 2^2 + \ldots + 6 \times 6^2\right] \\ &= \frac{1}{36} \times 1,911 \\ &= 53.08333 \qquad = 3.5 \times \frac{91}{6} = E\{X\} \times E\{Y^2\} \end{aligned}$$

Addition rule

$$E\{X_1 + X_2 + \ldots + X_n\} \;=\; E\{X_1\} + E\{X_2\} + \ldots + E\{X_n\}$$

- It suffices to prove this for $n = 2$.

$$\begin{aligned}
&E\{X+Y\} \\
&\quad = \sum_x \sum_y (x+y)\, P\{X=x, Y=y\} \\
&\quad = \sum_x \sum_y x\, P\{X=x, Y=y\} + \sum_x \sum_y y\, P\{X=x, Y=y\} \\
&\quad = \sum_x x \sum_y P\{X=x, Y=y\} + \sum_y y \sum_x P\{X=x, Y=y\} \\
&\quad = \sum_x x\, P\{X=x\} + \sum_y y\, P\{Y=y\} \\
&\quad = E\{X\} + E\{Y\}
\end{aligned}$$

Multiplication rule

$$\boxed{E\{XY\} = \sum_{(x,y)} xyP\{X = x, Y = y\}}$$

- If X and Y are <u>independent</u>

$$\begin{aligned}
E\{XY\} &= \sum_{(x,y)} xy\, P\{X = x, Y = y\} \\
&= \sum_{(x,y)} xy\, P\{X = x\} \times P\{Y = y\} \\
&= \left(\sum_{x} x\, P\{X = x\}\right) \times \left(\sum_{y} yP\{Y = y\}\right) \\
&= E\{X\}\, E\{Y\}
\end{aligned}$$

Back to the Binomial: $E\{X^2\}$.

- $X \sim \text{Bin}(n,p) \;\Rightarrow\; X = I_1 + I_2 + \ldots + I_n.$
- $E\{I_k\} = p, \quad E\{I_k^2\} = E\{I_k\} = p.$

$$
\begin{aligned}
E\left\{X^2\right\} &= E\left\{(I_1 + I_2 + \ldots + I_n)^2\right\} \\
&= E\left\{I_1^2 + \ldots + I_n^2 + \sum_{i \neq j} I_i I_j\right\} \\
&= \sum_{i=1}^{n} E\left\{I_i^2\right\} \;+\; \sum_{i \neq j} E\left\{I_i I_j\right\} \\
&= \sum_{i=1}^{n} E\left\{I_i^2\right\} \;+\; \sum_{i \neq j} E\{I_i\} \times E\{I_j\} \\
&= np \;+\; n(n-1)p^2 \quad = \; np(1-p) \;+ n^2 p^2
\end{aligned}
$$

Expectation and Prediction

- AIM: To predict a value of X.
- LOSS FUNCTION: $L(x, b)$ measures the "cost" of the guess b if the real value of X is x.
- POLICY: Minimise $E\{L(X, b)\}$.

1: "Right or wrong"

$$L(x, b) = \begin{cases} 0 & \text{if } x = b \ , \\ 1 & \text{otherwise.} \end{cases}$$

- $\Rightarrow E\{L(X, b)\} = P\{X \neq b\}$.
- Minimising this $\equiv$ <u>maximising</u> $1 - P\{X \neq b\} = P\{X = b\}$.
- $\Rightarrow$ choose $b =$ <u>mode</u> (any) of distribution of X.

2: "Squared error"

$$L(x, b) \;=\; (x - b)^2$$

$$E\{L(X, b)\} \;=\; E\{(X - b)^2\} \;=\; E\{X^2\} - 2bE\{X\} + b^2$$

Therefore

$$\frac{d}{db} E\{L(x, b)\} \;=\; -2E\{X\} + 2b$$

$\Rightarrow \quad b = E\{X\}$ gives the <u>unique</u> best predictor for X under squared error loss.

3: "Absolute error"

$$L(x, b) \;=\; |x - b|$$

- Under "absolute error" a median is the best guess.

  See Lecture 22 for this in the continuous case.