Chapter 7 part 1

Properties of Expectations

Jingchen (Monika) Hu

Vassar College

MATH 241

Outline

Expected value of $g(X, Y)$

Recap: expectation of random variable $g(X)$

- Discrete case $E[g(X)] = \sum_{\text{all } x} g(x) f(x)$
- Continuous case $E[g(X)] = \int_{-\infty}^{\infty} g(x) f(x) dx$

Expected value of $g(X, Y)$

Recap: expectation of random variable $g(X)$

- Discrete case $E[g(X)] = \sum_{\text{all } x} g(x)f(x)$
- Continuous case $E[g(X)] = \int_{-\infty}^{\infty} g(x)f(x)dx$

Suppose $g(X, Y)$ is a real-valued function of random variables X and Y, then

- Discrete case

$$E[g(X, Y)] = \sum_{\text{all } x} \sum_{\text{all } y} g(x, y)f(x, y)$$

- Continuous case

$$E[g(X, Y)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x, y)f(x, y)dxdy$$

Example: let X and Y be random variables with joint pdf $f(x, y)$. Find $E(X + Y)$

Example: let X and Y be random variables with joint pdf $f(x, y)$. Find $E(X+Y)$

$$E(X+Y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x+y)f(x,y)\ dx\ dy$$

Example: let X and Y be random variables with joint pdf $f(x, y)$. Find $E(X + Y)$

$$
\begin{aligned}
E(X+Y) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x+y)f(x,y)\ dx\ dy \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xf(x,y)\ {\color{red}dy\ dx}
\end{aligned}
$$

Example: let X and Y be random variables with joint pdf $f(x, y)$. Find $E(X+Y)$

$$
\begin{aligned}
E(X+Y) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x+y)f(x,y)\ dx\ dy \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xf(x,y)\ dy\ dx + \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} yf(x,y)\ dx\ dy
\end{aligned}
$$

Example: let X and Y be random variables with joint pdf $f(x, y)$. Find $E(X+Y)$

$$\begin{aligned} E(X+Y) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x+y)f(x,y)\ dx\ dy \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xf(x,y)\ {\color{red}dy\ dx} + \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} yf(x,y)\ dx\ dy \\ &= \int_{-\infty}^{\infty} x\left[\int_{-\infty}^{\infty} f(x,y)dy\right]dx \end{aligned}$$

Example: let X and Y be random variables with joint pdf $f(x,y)$. Find $E(X+Y)$

$$\begin{aligned} E(X+Y) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x+y)f(x,y)\ dx\ dy \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}xf(x,y)\ {\color{red}dy\ dx} + \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}yf(x,y)\ dx\ dy \\ &= \int_{-\infty}^{\infty}x\left[\int_{-\infty}^{\infty}f(x,y)dy\right]dx + \int_{-\infty}^{\infty}y\left[\int_{-\infty}^{\infty}f(x,y)dx\right]dy \end{aligned}$$

Example: let X and Y be random variables with joint pdf $f(x,y)$. Find $E(X+Y)$

$$
\begin{aligned}
E(X+Y) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x+y)f(x,y)\ dx\ dy \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xf(x,y)\ dy\ dx + \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} yf(x,y)\ dx\ dy \\
&= \int_{-\infty}^{\infty} x\left[\int_{-\infty}^{\infty} f(x,y)dy\right]dx + \int_{-\infty}^{\infty} y\left[\int_{-\infty}^{\infty} f(x,y)dx\right]dy \\
&= \int_{-\infty}^{\infty} xf_X(x)dx + \int_{-\infty}^{\infty} yf_Y(y)dy
\end{aligned}
$$

Example: let X and Y be random variables with joint pdf $f(x, y)$. Find $E(X+Y)$

$$
\begin{aligned}
E(X+Y) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x+y)f(x,y)\ dx\ dy \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}xf(x,y)\ dy\ dx + \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}yf(x,y)\ dx\ dy \\
&= \int_{-\infty}^{\infty}x\left[\int_{-\infty}^{\infty}f(x,y)dy\right]dx + \int_{-\infty}^{\infty}y\left[\int_{-\infty}^{\infty}f(x,y)dx\right]dy \\
&= \int_{-\infty}^{\infty}xf_X(x)dx + \int_{-\infty}^{\infty}yf_Y(y)dy \\
&= E(X) + E(Y)
\end{aligned}
$$

Expectation of sums of two random variables

$$E(X+Y) = E(X) + E(Y)$$

Expectation of sums of two random variables

$$E(X+Y) = E(X) + E(Y)$$

- It's not difficult to show that if either (or both) of the X, Y is discrete, this formula still holds.

Expectation of sums of two random variables

$$E(X+Y) = E(X) + E(Y)$$

- It's not difficult to show that if either (or both) of the X, Y is discrete, this formula still holds.
- This results does not require X and Y to be independent.

Expectation of sums of two random variables

$$E(X+Y) = E(X) + E(Y)$$

- It's not difficult to show that if either (or both) of the X, Y is discrete, this formula still holds.
- This results does not require X and Y to be independent.
- This can be generalized to n random variables

$$E(X_1 + X_2 + \cdots + X_n) = E(X_1) + E(X_2) + \cdots + E(X_n)$$

Expectation of sums of two random variables

$$E(X+Y) = E(X) + E(Y)$$

- It's not difficult to show that if either (or both) of the X, Y is discrete, this formula still holds.
- This results does not require X and Y to be independent.
- This can be generalized to n random variables

$$E(X_1 + X_2 + \cdots + X_n) = E(X_1) + E(X_2) + \cdots + E(X_n)$$

- How about $E(XY)$?

Let X and Y be INDEPENDENT random variables with joint pdf $f(x, y)$. Find $E(XY)$.

Let X and Y be INDEPENDENT random variables with joint pdf $f(x, y)$. Find $E(XY)$.

$$E(XY) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xyf(x, y)\ dx\ dy$$

Let X and Y be INDEPENDENT random variables with joint pdf $f(x,y)$. Find $E(XY)$.

$$
\begin{aligned}
E(XY) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xyf(x,y)\;dx\;dy \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xyf_X(x)f_Y(y)\;dx\;dy
\end{aligned}
$$

Let X and Y be INDEPENDENT random variables with joint pdf $f(x, y)$. Find $E(XY)$.

$$
\begin{aligned}
E(XY) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xyf(x,y)\ dx\ dy \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xyf_X(x)f_Y(y)\ dx\ dy \\
&= \int_{-\infty}^{\infty} yf_Y(y)\left[\int_{-\infty}^{\infty} xf_X(x)dx\right] dy
\end{aligned}
$$

Let X and Y be INDEPENDENT random variables with joint pdf $f(x,y)$. Find $E(XY)$.

$$
\begin{aligned}
E(XY) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xyf(x,y)\ dx\ dy \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xyf_X(x)f_Y(y)\ dx\ dy \\
&= \int_{-\infty}^{\infty} yf_Y(y)\left[\int_{-\infty}^{\infty} xf_X(x)dx\right] dy \\
&= E(X)\int_{-\infty}^{\infty} yf_Y(y)dy
\end{aligned}
$$

Let X and Y be INDEPENDENT random variables with joint pdf $f(x,y)$. Find $E(XY)$.

$$
\begin{aligned}
E(XY) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xyf(x,y)\;dx\;dy \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xyf_X(x)f_Y(y)\;dx\;dy \\
&= \int_{-\infty}^{\infty} yf_Y(y)\left[\int_{-\infty}^{\infty} xf_X(x)dx\right]dy \\
&= E(X)\int_{-\infty}^{\infty} yf_Y(y)dy \\
&= E(X)E(Y)
\end{aligned}
$$

Let X and Y be INDEPENDENT random variables with joint pdf $f(x, y)$. Find $E(XY)$.

$$
\begin{aligned}
E(XY) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xyf(x,y)\ dx\ dy \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xyf_X(x)f_Y(y)\ dx\ dy \\
&= \int_{-\infty}^{\infty} yf_Y(y)\left[\int_{-\infty}^{\infty} xf_X(x)dx\right] dy \\
&= E(X)\int_{-\infty}^{\infty} yf_Y(y)dy \\
&= E(X)E(Y)
\end{aligned}
$$

Note: (1) this formula only holds when X and Y are independent.
(2) This is not a sufficient condition for independence.

Recap

Expectation of sum

$$E[X_1 + X_2 + \cdots + X_n] = E[X_1] + E[X_2] + \cdots + E[X_n]$$

Recap

Expectation of sum

$$E[X_1 + X_2 + \cdots + X_n] = E[X_1] + E[X_2] + \cdots + E[X_n]$$

$X_1, X_2, \ldots, X_n$ are independent $\Longrightarrow$ $\not\Longleftarrow$

$$E[X_1 X_2 \cdots X_n] = E[X_1]E[X_2] \cdots E[X_n]$$

Outline

Covariance

Definition

Covariance of two random variables X and Y is defined as

$$Cov(X,Y) = E[(X - E[X])(Y - E[Y])]$$

Covariance

Definition

Covariance of two random variables X and Y is defined as

$$Cov(X,Y) = E[(X - E[X])(Y - E[Y])]$$

- Simplification

$$\begin{aligned} Cov(X,Y) &= E[(X - \mu_X)(Y - \mu_Y)] \\ &= E[XY + \mu_X \mu_Y - X\mu_Y - Y\mu_X]) \\ &= E[XY] - \mu_X \mu_Y \end{aligned}$$

Covariance

Definition

Covariance of two random variables X and Y is defined as

$$Cov(X, Y) = E[(X - E[X])(Y - E[Y])]$$

- Simplification

$$\begin{aligned} Cov(X, Y) &= E[(X - \mu_X)(Y - \mu_Y)] \\ &= E[XY + \mu_X \mu_Y - X\mu_Y - Y\mu_X]) \\ &= E[XY] - \mu_X \mu_Y \end{aligned}$$

- Recall

$$\begin{aligned} E[XY] &= \int \int xy\ f(x, y)\ dx\ dy \quad \text{if continuous} \\ &= \sum_x \sum_y xy\ f(x, y) \quad \text{if discrete} \end{aligned}$$

Properties of $Cov(X, Y) = E[XY] - E[X]E[Y]$

- $Cov(X, Y) = Cov(Y, X)$

Properties of $Cov(X, Y) = E[XY] - E[X]E[Y]$

- $Cov(X, Y) = Cov(Y, X)$
- $Cov(X, c) = 0$

Properties of $Cov(X, Y) = E[XY] - E[X]E[Y]$

- $Cov(X, Y) = Cov(Y, X)$
- $Cov(X, c) = 0$
- $Cov(X, X) = Var(X)$

Properties of $Cov(X, Y) = E[XY] - E[X]E[Y]$

- $Cov(X, Y) = Cov(Y, X)$
- $Cov(X, c) = 0$
- $Cov(X, X) = Var(X)$
- $Cov(aX, bY) = ab\ Cov(X, Y)$

Properties of $Cov(X, Y) = E[XY] - E[X]E[Y]$

- $Cov(X, Y) = Cov(Y, X)$
- $Cov(X, c) = 0$
- $Cov(X, X) = Var(X)$
- $Cov(aX, bY) = ab\ Cov(X, Y)$

$$\begin{aligned} Cov(aX, bY) &= E[abXY] - E[aX]E[bY] \\ &= ab\ E[XY] - ab\ E[X]E[Y] \end{aligned}$$

Properties of $Cov(X, Y) = E[XY] - E[X]E[Y]$

- $Cov(X, Y) = Cov(Y, X)$
- $Cov(X, c) = 0$
- $Cov(X, X) = Var(X)$
- $Cov(aX, bY) = ab\ Cov(X, Y)$

$$\begin{aligned} Cov(aX, bY) &= E[abXY] - E[aX]E[bY] \\ &= ab\ E[XY] - ab\ E[X]E[Y] \end{aligned}$$

- $Cov(X + a, Y + b) = Cov(X, Y)$

Properties of $Cov(X, Y) = E[XY] - E[X]E[Y]$

- $Cov(X, Y) = Cov(Y, X)$
- $Cov(X, c) = 0$
- $Cov(X, X) = Var(X)$
- $Cov(aX, bY) = ab\ Cov(X, Y)$

$$\begin{aligned} Cov(aX, bY) &= E[abXY] - E[aX]E[bY] \\ &= ab\ E[XY] - ab\ E[X]E[Y] \end{aligned}$$

- $Cov(X + a, Y + b) = Cov(X, Y)$

$$\begin{aligned} Cov(X + a, Y + b) =& E[(X + a)(Y + b)] - E[X + a]E[Y + b] \\ =& E[XY + aY + bX + ab] \\ & - (E[X] + a)(E(Y) + b) \\ =& E[XY] + E[aY] + E[bX] + ab \\ & - E[X]E[Y] - a\ E[Y] - b\ E[X] - ab \\ =& E[XY] - E[X]E[Y] \end{aligned}$$

Covariance of sums of random variables

$$Cov\left(\sum_{i=1}^{n} X_i, \sum_{j=1}^{m} Y_j\right) = \sum_{i=1}^{n}\sum_{j=1}^{m} Cov\left(X_i, Y_j\right)$$

Covariance of sums of random variables

$$Cov\left(\sum_{i=1}^{n} X_i, \sum_{j=1}^{m} Y_j\right) = \sum_{i=1}^{n}\sum_{j=1}^{m} Cov\left(X_i, Y_j\right)$$

A special case

$$Var\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} Var(X_i) + 2\sum_{1\leq i<j\leq n} Cov(X_i, X_j)$$

Covariance of sums of random variables

$$Cov\left(\sum_{i=1}^{n} X_i, \sum_{j=1}^{m} Y_j\right) = \sum_{i=1}^{n}\sum_{j=1}^{m} Cov\left(X_i, Y_j\right)$$

A special case

$$Var\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} Var(X_i) + 2 \sum_{1\leq i<j\leq n} Cov(X_i, X_j)$$

Some more special cases

$$Var(X+Y) = Var(X) + Var(Y) + 2Cov(X,Y)$$
$$Var(X-Y) = Var(X) + Var(Y) - 2Cov(X,Y)$$

Zero covariance and independence

- X and Y are independent $\Longrightarrow Cov(X,Y) = 0$

Zero covariance and independence

- X and Y are independent $\Longrightarrow Cov(X,Y) = 0$

$$Cov(X,Y) = E[XY] - E[X]E[Y] = E[X]E[Y] - E[X]E[Y] = 0$$

Zero covariance and independence

- X and Y are independent $\Longrightarrow Cov(X,Y) = 0$

$$Cov(X,Y) = E[XY] - E[X]E[Y] = E[X]E[Y] - E[X]E[Y] = 0$$

- $X_1, X_2, \ldots, X_n$ are independent $\Longrightarrow$

$$Var\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} Var(X_i)$$

Zero covariance and independence

- X and Y are independent $\implies Cov(X,Y) = 0$

$$Cov(X,Y) = E[XY] - E[X]E[Y] = E[X]E[Y] - E[X]E[Y] = 0$$

- $X_1, X_2, \ldots, X_n$ are independent $\implies$

$$Var\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} Var(X_i)$$

- $Cov(X,Y) = 0 \not\Longrightarrow X$ and Y are independent
  Counter example?

Counter example

Let $X \sim \text{Unif}(-0.5, 0.5)$ and $Y = X^2$. Find $Cov(X, Y)$, and decide if X and Y and independent.

Counter example

Let $X \sim$ Unif$(-0.5, 0.5)$ and $Y = X^2$. Find $Cov(X, Y)$, and decide if X and Y and independent.

Covariance

Counter example

Let $X \sim$ Unif$(-0.5, 0.5)$ and $Y = X^2$. Find $Cov(X, Y)$, and decide if X and Y and independent.

Covariance

$$E[X] = 0, \; E[XY] = E[X^3] = \int_{-0.5}^{0.5} x^3 dx = \frac{x^4}{4}\Big|_{-0.5}^{0.5} = 0$$

Counter example

Let $X \sim$ Unif$(-0.5, 0.5)$ and $Y = X^2$. Find $Cov(X, Y)$, and decide if X and Y and independent.

Covariance

$$E[X] = 0,\ E[XY] = E[X^3] = \int_{-0.5}^{0.5} x^3 dx = \frac{x^4}{4}\Big|_{-0.5}^{0.5} = 0$$

$$Cov(X, Y) = E[XY] - E[X]E[Y] = E[X^3] - E[X]E[X^2] = 0$$

Counter example

Let $X \sim$ Unif$(-0.5, 0.5)$ and $Y = X^2$. Find $Cov(X, Y)$, and decide if X and Y and independent.

Covariance

$$E[X] = 0,\ E[XY] = E[X^3] = \int_{-0.5}^{0.5} x^3 dx = \frac{x^4}{4}\Big|_{-0.5}^{0.5} = 0$$

$$Cov(X, Y) = E[XY] - E[X]E[Y] = E[X^3] - E[X]E[X^2] = 0$$

Independence:

Counter example

Let $X \sim$ Unif$(-0.5, 0.5)$ and $Y = X^2$. Find $Cov(X, Y)$, and decide if X and Y and independent.

Covariance

$$E[X] = 0,\ E[XY] = E[X^3] = \int_{-0.5}^{0.5} x^3 dx = \frac{x^4}{4}\Big|_{-0.5}^{0.5} = 0$$

$$Cov(X, Y) = E[XY] - E[X]E[Y] = E[X^3] - E[X]E[X^2] = 0$$

Independence: since Y depends on X, so not independent.

Counter example

Let $X \sim$ Unif$(-0.5, 0.5)$ and $Y = X^2$. Find $Cov(X, Y)$, and decide if X and Y and independent.

Covariance

$$E[X] = 0, \; E[XY] = E[X^3] = \int_{-0.5}^{0.5} x^3 dx = \frac{x^4}{4}\Big|_{-0.5}^{0.5} = 0$$

$$Cov(X, Y) = E[XY] - E[X]E[Y] = E[X^3] - E[X]E[X^2] = 0$$

Independence: since Y depends on X, so not independent.
(To be more rigorous, we need to show

$$f(x, y) \neq f_X(x) f_Y(y)$$

for some $x, y \in \mathbb{R}$.)

Correlation

Since $Cov(X, Y)$ depends on the magnitude of X and Y we would prefer to have a measure of association that is not effected by arbitrary changes in the scales of the random variables.

Definition

The most common measure of linear association is correlation which is defined as

$$\rho(X, Y) = \frac{Cov(X, Y)}{\sqrt{Var(X)}\sqrt{Var(Y)}}$$

Correlation

Since $Cov(X, Y)$ depends on the magnitude of X and Y we would prefer to have a measure of association that is not effected by arbitrary changes in the scales of the random variables.

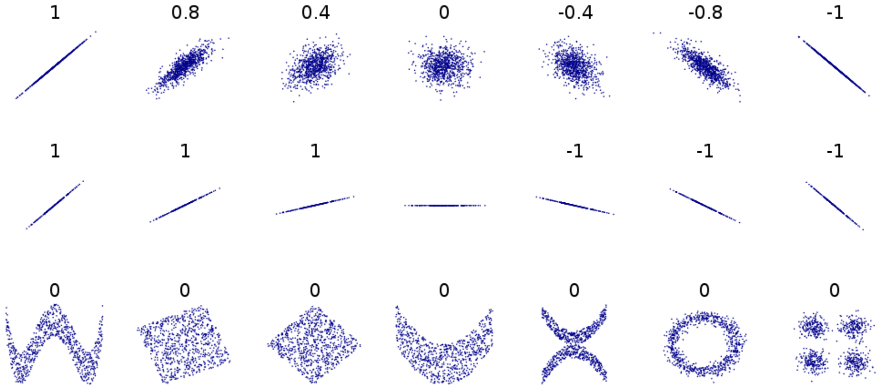
Definition

The most common measure of linear association is correlation which is defined as

$$\rho(X, Y) = \frac{Cov(X, Y)}{\sqrt{Var(X)}\,\sqrt{Var(Y)}}$$

- range: $-1 \leq \rho(X, Y) \leq 1$
- the magnitude (i.e. absolute value) of the $\rho(X, Y)$ measures the strength of the linear association
- the sign determines if it is a positive or negative relationship.
- if $\rho(X, Y) = 0$, then X and Y are said to be uncorrelated.

Correlation



Outline

Conditional expectation

- The discrete case: for all $p_Y(y) > 0$
  - Conditional pmf: $p_{X|Y}(x \mid y) = P\{X = x \mid Y = y\} = \frac{p(x,y)}{p_Y(y)}$

Definition

The conditional expectation of X given that $Y = y$ is

$$E[X \mid Y = y] = \sum_x xP\{X = x \mid Y = y\} = \sum_x xp_{X|Y}(x \mid y)$$

Conditional expectation

- The discrete case: for all $p_Y(y) > 0$
  - Conditional pmf: $p_{X|Y}(x \mid y) = P\{X = x \mid Y = y\} = \frac{p(x,y)}{p_Y(y)}$

Definition

The conditional expectation of X given that $Y = y$ is

$$E[X \mid Y = y] = \sum_x xP\{X = x \mid Y = y\} = \sum_x xp_{X|Y}(x \mid y)$$

- The continuous case: for all $f_Y(y) > 0$
  - Conditional pdf: $f_{X|Y}(x \mid y) = \frac{f(x,y)}{f_Y(y)}$

Definition

The conditional expectation of X given that $Y = y$ is

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

Properties of expectations remain

- Expectation of a function of a random variable
  - The discrete case:

$$E[g(X) \mid Y = y] = \sum_x g(x) p_{X|Y}(x \mid y)$$

  - The continuous case:

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

Properties of expectations remain

- Expectation of a function of a random variable
  - The discrete case:

$$E[g(X) \mid Y = y] = \sum_{x} g(x) p_{X|Y}(x \mid y)$$

  - The continuous case:

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

- Expectation of sum of random variables

$$E[\sum_{i=1}^{n} X_i \mid Y = y] = \sum_{i=1}^{n} E[X_i \mid Y = y]$$

Properties of expectations remain

- Expectation of a function of a random variable
  - The discrete case:

$$E[g(X) \mid Y = y] = \sum_x g(x) p_{X|Y}(x \mid y)$$

  - The continuous case:

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

- Expectation of sum of random variables

$$E[\sum_{i=1}^{n} X_i \mid Y = y] = \sum_{i=1}^{n} E[X_i \mid Y = y]$$

- Think about the condition $Y = y$ as taking expectation of X on a reduced sample space consisting only of outcomes for which $Y = y$

Computing expectations by conditioning

$$E[X] = E[E[X \mid Y]]$$

- A very important property of conditional expectation
- Think of $E[X|Y]$ as a random variable (when $Y = y$)
- The discrete case:

$$E[X] = \sum_y E[X \mid Y = y] P\{Y = y\}$$

- Intuition:
  - $E[E[X \mid Y]]$ is a weighted average of $E[X \mid Y]$, where weights are $P\{Y = y\}$ (the probability of the condition)
  - Similar to the "law of total probability" $P(E) = \sum_{i=1}^{n} P(E \mid F_i) P(F_i)$

Computing expectations by conditioning

$$E[X] = E[E[X \mid Y]]$$

- A very important property of conditional expectation
- Think of $E[X|Y]$ as a random variable (when $Y = y$)
- The discrete case:

$$E[X] = \sum_y E[X \mid Y = y] P\{Y = y\}$$

- Intuition:
  - $E[E[X \mid Y]]$ is a weighted average of $E[X \mid Y]$, where weights are $P\{Y = y\}$ (the probability of the condition)
  - Similar to the "law of total probability" $P(E) = \sum_{i=1}^{n} P(E \mid F_i) P(F_i)$
- The continuous case:

$$E[X] = \int_{-\infty}^{\infty} E[X \mid Y = y] f_Y(y) dy$$

A miner is trapped in a mine containing 3 doors. The 1st door leads to a tunnel that will take him to safety after 3 hours of travel. The 2nd door leads to a tunnel that will return him to the mine after 5 hours of travel. The 3rd door leads to a tunnel that will return him to the mine after 7 hours. If we assume that the miner is at all times equally likely to choose any one of the doors, what is the expected length of time until he reaches safety?

A miner is trapped in a mine containing 3 doors. The 1st door leads to a tunnel that will take him to safety after 3 hours of travel. The 2nd door leads to a tunnel that will return him to the mine after 5 hours of travel. The 3rd door leads to a tunnel that will return him to the mine after 7 hours. If we assume that the miner is at all times equally likely to choose any one of the doors, what is the expected length of time until he reaches safety?

- Let X denote the amount of time in hours until he reaches safety, and Y denote the door he initially chooses.

$$\begin{aligned} E[X] &= E[X \mid Y=1]P\{Y=1\} + E[X \mid Y=2]P\{Y=2\} + E[X \mid Y=3]P\{Y=3\} \\ &= \frac{1}{3}(E[X \mid Y=1] + E[X \mid Y=2] + E[X \mid Y=3]) \end{aligned}$$

A miner is trapped in a mine containing 3 doors. The 1st door leads to a tunnel that will take him to safety after 3 hours of travel. The 2nd door leads to a tunnel that will return him to the mine after 5 hours of travel. The 3rd door leads to a tunnel that will return him to the mine after 7 hours. If we assume that the miner is at all times equally likely to choose any one of the doors, what is the expected length of time until he reaches safety?

- Let X denote the amount of time in hours until he reaches safety, and Y denote the door he initially chooses.

$$\begin{aligned} E[X] &= E[X \mid Y=1]P\{Y=1\} + E[X \mid Y=2]P\{Y=2\} + E[X \mid Y=3]P\{Y=3\} \\ &= \frac{1}{3}(E[X \mid Y=1] + E[X \mid Y=2] + E[X \mid Y=3]) \end{aligned}$$

- Note that
$E[X|Y=1] = 3, E[X \mid Y=2] = 5 + E[X], E[X \mid Y=3] = 7 + E[X]$
- Therefore $E[X] = \frac{1}{3}(3 + 5 + E[X] + 7 + E[X])$, which gives $E[X] = 15$

Computing probabilities by conditioning

- Let E denote an arbitrary event, and define the indicator random variable X as
$$X = \begin{cases} 1 & \text{if } E \text{ occurs} \\ 0 & \text{if } E \text{ does not occur} \end{cases}$$
- Then $E[X] = P(E), E[X \mid Y = y] = P(E \mid Y = y)$ for any random variable Y
- The discrete case:
$$P(E) = \sum_y P(E \mid Y = y)p(Y = y)$$
related to $P(E) = \sum_{i=1}^n P(E \mid F_i)P(F_i)$

Computing probabilities by conditioning

- Let E denote an arbitrary event, and define the indicator random variable X as
$$X = \begin{cases} 1 & \text{if } E \text{ occurs} \\ 0 & \text{if } E \text{ does not occur} \end{cases}$$
- Then $E[X] = P(E), E[X \mid Y = y] = P(E \mid Y = y)$ for any random variable Y
- The discrete case:
$$P(E) = \sum_y P(E \mid Y = y) p(Y = y)$$
related to $P(E) = \sum_{i=1}^n P(E \mid F_i) P(F_i)$
- The continuous case:
$$P(E) = \int_{-\infty}^{\infty} P(E \mid Y = y) f_Y(y) dy$$

Conditional variance

- Similarly to the conditional expectation, we can define the conditional variance of X given that $Y = y$

Definition

$$Var(X \mid Y = y) = E[(X - E[X \mid Y = y])^2 \mid Y = y]$$

Conditional variance

- Similarly to the conditional expectation, we can define the conditional variance of X given that $Y = y$

Definition

$$Var(X \mid Y = y) = E[(X - E[X \mid Y = y])^2 \mid Y = y]$$

- A very useful conditional variance formula

$$Var(X) = E[Var(X \mid Y)] + Var(E[X \mid Y])$$