STSCI 5080
Probability Models and Inference
Lecture 12: Expected Values

October 4, 2018

Standard deviation

Definition

The standard deviation of a random variable X is defined by

$$\sqrt{\mathrm{Var}(X)}$$

provided that $E(X^2) < \infty$.

Covariance

Definition

The covariance between two random variables X and Y is defined by

$$\mathrm{Cov}(X, Y) = E[\{X - E(X)\}\{Y - E(Y)\}]$$

provided that $E(X^2) < \infty$ and $E(Y^2) < \infty$.

- The covariance is a measure on dependence of two variables.
- Can be negative and positive.

Properties of covariance

Theorem

(a) $|\mathrm{Cov}(X, Y)| \leq \sqrt{\mathrm{Var}(X)}\sqrt{\mathrm{Var}(Y)}$.

(b) $\mathrm{Cov}(X, X) = \mathrm{Var}(X)$ *and* $\mathrm{Cov}(X, -X) = -\mathrm{Var}(X)$.

(c) $\mathrm{Cov}(X, Y) = E(XY) - E(X)E(Y)$. *In particular, if* X *and* Y *are independent, then* $\mathrm{Cov}(X, Y) = 0$.

Proof

(a) By the Cauchy-Schwarz inequality,

$$\begin{aligned}
|\mathrm{Cov}(X, Y)| &\le E[|\{X - E(X)\}\{Y - E(Y)\}|] \\
&\le \sqrt{E[\{X - E(X)\}^2]}\sqrt{E[\{Y - E(Y)\}^2]} \\
&= \sqrt{\mathrm{Var}(X)}\sqrt{\mathrm{Var}(Y)}.
\end{aligned}$$

(b) Skip.

(c) Since

$$\{X - E(X)\}\{Y - E(Y)\} = XY - E(X)Y - E(Y)X + E(X)E(Y),$$

we have

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

$\mathrm{Cov} = 0 \not\Rightarrow$ independence

Example

A random vector (X, Y) takes $(1, 0), (0, 1), (-1, 0), (0, -1)$ with probability $1/4$ each. Then $\mathrm{Cov}(X, Y) = 0$ but X and Y are not independent.

Example 12.1

Example

Let (X, Y) be a uniform random vector on the triangle region $A = \{(x, y) \mid x + y \leq 1, x, y \geq 0\}$. Find $\mathrm{Cov}(X, Y)$.

Example 12.1

Example

Let (X, Y) be a uniform random vector on the triangle region $A = \{(x, y) \mid x + y \leq 1, x, y \geq 0\}$. Find $\mathrm{Cov}(X, Y)$.

The joint pdf is

$$f(x, y) = 2 \text{ if } (x, y) \in A$$

and $f(x, y) = 0$ elsewhere. So,

$$E(XY) = 2 \iint_A (xy)dxdy = 2 \int_0^1 x \left\{ \int_0^{1-x} ydy \right\} dx = \int_0^1 x(1-x)^2 dx = \frac{1}{12}.$$

On the other hand, the marginal pdf of X is $f_X(x) = 2(1-x)$ for $0 \leq x \leq 1$, and so

$$E(X) = 2 \int_0^1 x(1-x)dx = \frac{1}{3}.$$

By symmetry, $E(Y) = 1/3$, so that $\mathrm{Cov}(X, Y) = 1/12 - 1/9 = -1/36$.

Variance of linear combination of random variables

Theorem

$\mathrm{Var}(aX + bY) = a^2\mathrm{Var}(X) + 2ab\mathrm{Cov}(X, Y) + b^2\mathrm{Var}(Y)$. *In particular, if* X *and* Y *are independent, then* $\mathrm{Var}(X + Y) = \mathrm{Var}(X) + \mathrm{Var}(Y)$.

Proof.

Let $\widetilde{X} = X - E(X)$ and $\widetilde{Y} = Y - E(Y)$. Then

$$\begin{aligned}
\mathrm{Var}(aX + bY) &= E\{(a\widetilde{X} + b\widetilde{Y})^2\} = E(a^2\widetilde{X}^2 + 2ab\widetilde{X}\widetilde{Y} + b^2\widetilde{Y}^2) \\
&= a^2E(\widetilde{X}^2) + 2abE(\widetilde{X}\widetilde{Y}) + b^2E(\widetilde{Y}^2) \\
&= a^2\mathrm{Var}(X) + 2ab\mathrm{Cov}(X, Y) + b^2\mathrm{Var}(Y).
\end{aligned}$$

□

Theorem

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

In particular, if $X_1, \ldots, X_n$ are independent, then

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

Example 12.2

Example

If $Y \sim Bin(n, p)$, then find $\text{Var}(Y)$.

Example 12.2

Example

If $Y \sim Bin(n, p)$, then find $\mathrm{Var}(Y)$.

By definition, $Y = X_1 + \cdots + X_n$ for independent Bernoulli trials $X_1, \ldots, X_n$ with success probability p. In addition, each X_i has variance $p(1-p)$:

$$\mathrm{Var}(X_i) = p(1-p).$$

Hence, we have

$$\mathrm{Var}(Y) = \mathrm{Var}(X_1) + \cdots + \mathrm{Var}(X_n) = np(1-p).$$

Correlation

- The covariance depends on the unit (e.g. kg vs. lbs) to measure X or Y.
- "Dependence" should be independent of the unit.

Definition

The correlation between X and Y is defined by

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

provided that $E(X^2) < \infty, E(Y^2) < \infty, \mathrm{Var}(X) > 0$, and $\mathrm{Var}(Y) > 0$.

Properties of correlation

Theorem

(a) *The correlation is independent of the unit to measure* X *or* Y*:* $\mathrm{Corr}(aX, bY) = \mathrm{Corr}(X, Y)$ *for any* $a, b > 0$*.*

(b) $|\mathrm{Corr}(X, Y)| \leq 1$*.*

Proof?

Theorem

Suppose that the correlation $\mathrm{Corr}(X, Y)$ *is defined and let* $\mu_X = E(X), \mu_Y = E(Y), \sigma_X^2 = \mathrm{Var}(X)$, *and* $\sigma_Y^2 = \mathrm{Var}(Y)$.

(a) *If* $\mathrm{Corr}(X, Y) = 1$, *then*

$$X = aY + b$$

for some constants $a > 0$ *and* $-\infty < b < \infty$.

(b) *If* $\mathrm{Corr}(X, Y) = -1$, *then*

$$X = aY + b$$

for some constants $a < 0$ *and* $-\infty < b < \infty$.

Proof of Case (a)

Let $\widetilde{X} = (X - \mu_X)/\sigma_X$ and $\widetilde{Y} = (Y - \mu_Y)/\sigma_Y$, so that

$$E(\widetilde{X}) = E(\widetilde{Y}) = 0 \quad \text{and} \quad \mathrm{Var}(\widetilde{X}) = \mathrm{Var}(\widetilde{Y}) = 1.$$

Then

$$\begin{aligned} E\{(\widetilde{X} - \widetilde{Y})^2\} &= \mathrm{Var}(\widetilde{X} - \widetilde{Y}) = \underbrace{\mathrm{Var}(\widetilde{X})}_{=1} - 2\,\underbrace{\mathrm{Cov}(\widetilde{X}, \widetilde{Y})}_{=\mathrm{Corr}(X,Y)=1} + \underbrace{\mathrm{Var}(\widetilde{Y})}_{=1} \\ &= 1 - 2 + 1 = 0. \end{aligned}$$

Hence, we have $\widetilde{X} = \widetilde{Y}$, i.e.,

$$X = \underbrace{\frac{\sigma_X}{\sigma_Y}}_{=a} Y + \underbrace{\mu_X - \frac{\sigma_X}{\sigma_Y}\mu_Y}_{=b}.$$

Example 12.3

Example

Let (X, Y) be a uniform random vector on the triangle region $A = \{(x, y) \mid x + y \leq 1, x, y \geq 0\}$. Find $\mathrm{Corr}(X, Y)$.

Example 12.3

Example

Let (X, Y) be a uniform random vector on the triangle region $A = \{(x, y) \mid x + y \leq 1, x, y \geq 0\}$. Find $\mathrm{Corr}(X, Y)$.

The covariance is

$$\mathrm{Cov}(X, Y) = -\frac{1}{36}.$$

The marginal pdf of X is $f_X(x) = 2(1 - x)$ for $0 \leq x \leq 1$, and so

$$E(X^2) = 2\int_0^1 x^2(1 - x)dx = \frac{1}{6}.$$

Since $E(X) = 1/3$, we have

$$\mathrm{Var}(X) = E(X^2) - \{E(X)\}^2 = \frac{1}{6} - \frac{1}{9} = \frac{1}{18}.$$

By symmetry, $\mathrm{Var}(Y) = 1/18$, so that

$$\mathrm{Corr}(X, Y) = \frac{-\frac{1}{36}}{\frac{1}{18}} = -\frac{1}{2}.$$

Conditional expectation

Definition

(a) If (X, Y) is discrete with joint pmf $p(x, y)$, then the conditional expectation of X given Y is defined by

$$E(X \mid Y = y) = \sum_x x p_{X|Y}(x \mid y) \quad \text{for any } y,$$

provided that $\sum_x |x| p_X(x) < \infty$.

(b) If (X, Y) is continuous with joint pdf $f(x, y)$, then the conditional expectation of X given Y is defined by

$$E(X \mid Y = y) = \int_{-\infty}^{\infty} x f_{X|Y}(x \mid y) dx \quad \text{for any } y,$$

provided that $\int_{-\infty}^{\infty} |x| f_X(x) dx < \infty$.

The conditional expectation $E(X \mid Y = y)$ is a function of y.

Theorem

(a) *If (X, Y) is discrete, then the conditional expectation of $g(X)$ given Y is given by*

$$E\{g(X) \mid Y = y\} = \sum_x g(x) p_{X|Y}(x \mid y) \quad \text{for any } y,$$

provided that $\sum_x |g(x)| p_X(x) < \infty$.

(b) *If (X, Y) is continuous with joint pdf $f(x, y)$, then the conditional expectation of $g(X)$ given Y is given by*

$$E\{g(X) \mid Y = y\} = \int_{-\infty}^{\infty} g(x) f_{X|Y}(x \mid y) dx \quad \text{for any } y,$$

provided that $\int_{-\infty}^{\infty} |g(x)| f_X(x) dx < \infty$.

Example 12.4

Example

Let (X, Y) be a uniform random vector on the triangle region $A = \{(x, y) \mid x + y \leq 1, x, y \geq 0\}$. Find the conditional expectation of X.

Example 12.4

Example

Let (X, Y) be a uniform random vector on the triangle region $A = \{(x, y) \mid x + y \leq 1, x, y \geq 0\}$. Find the conditional expectation of X.

Since the marginal pdf of Y is $f_Y(y) = 2(1 - y)$ for $0 \leq y \leq 1$, we have

$$f_{X|Y}(x \mid y) = \frac{1}{1 - y}$$

for $0 \leq x \leq 1 - y$ and $0 \leq y < 1$. Hence, we have

$$E(X \mid Y = y) = \frac{1}{1 - y} \int_0^{1-y} x dx = \frac{(1 - y)^2}{2(1 - y)} = \frac{1 - y}{2}$$

for $0 \leq y < 1$, and $E(X \mid Y = y) = 0$ elsewhere, so that $E(X \mid Y) = (1 - Y)/2$. (Why can we ignore the possibility that $Y < 0$ or $Y \geq 1$?)

Law of total expectation

Theorem

Let $E(X \mid Y) = E(X \mid Y = y)|_{y=Y}$ *(which is a random variable). Then*

$$E\{E(X \mid Y)\} = E(X).$$

Proof for the discrete case

We note that

$$E\{E(X \mid Y)\} = \sum_y E(X \mid Y = y) p_Y(y).$$

Since

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

we have

$$\begin{aligned} E\{E(X \mid Y)\} &= \sum_y \sum_x x \underbrace{p_{X|Y}(x \mid y) p_Y(y)}_{=p(x,y)} \\ &= \sum_x x \underbrace{\sum_y p(x, y)}_{p_X(x)} = E(X). \end{aligned}$$

Conditional variance

Definition

The conditional variance of X given Y is defined by

$$\mathrm{Var}(X \mid Y = y) = E(X^2 \mid Y = y) - \{E(X \mid Y = y)\}^2 \quad \text{for any } y,$$

provided that $E(X^2) < \infty$.

Theorem

Let $\mathrm{Var}(X \mid Y) = \mathrm{Var}(X \mid Y = y)|_{y=Y}$ *(which is a random variable). Then*

$$\mathrm{Var}(X) = \mathrm{Var}\{E(X \mid Y)\} + E\{\mathrm{Var}(X \mid Y)\}.$$

Proof

We note that

$$\begin{aligned}\mathrm{Var}\{E(X \mid Y)\} &= E[\{E(X \mid Y)\}^2] - [E\{E(X \mid Y)\}]^2 \\ &= E[\{E(X \mid Y)\}^2] - \{E(X)\}^2, \text{ and} \\ E\{\mathrm{Var}(X \mid Y)\} &= E\{E(X^2 \mid Y)\} - E[\{E(X \mid Y)\}^2] \\ &= E(X^2) - E[\{E(X \mid Y)\}^2].\end{aligned}$$

Hence, we have

$$\begin{aligned}&\mathrm{Var}\{E(X \mid Y)\} + E\{\mathrm{Var}(X \mid Y)\} \\ &= E[\{E(X \mid Y)\}^2] - \{E(X)\}^2 + E(X^2) - E[\{E(X \mid Y)\}^2] \\ &= E(X^2) - \{E(X)\}^2 = \mathrm{Var}(X).\end{aligned}$$

Compound Poisson random variable

Example

An insurance company receives a certain number of claims N per week that follows $Po(\lambda)$. The amount of each claim follows a cdf F. Then the amount of claims the company pays per week is

$$Y = \begin{cases} 0 & \text{if } N = 0 \\ \sum_{i=1}^{N} X_i & \text{if } N \geq 1 \end{cases},$$

where

- $N \sim Po(\lambda)$ (note: $E(N) = \mathrm{Var}(N) = \lambda$);
- For each n, $X_1, \ldots, X_n \sim F$ i.i.d.;
- For each n, N and $(X_1, \ldots, X_n)$ are independent.

The variable Y is called a compound Poisson random variable. Find $E(Y)$ and $\mathrm{Var}(Y)$.

Suppose that F is discrete. Let $p_n(x)$ denote the pmf of $\sum_{i=1}^n X_i$:

$$p_n(y) = P\left(\sum_{i=1}^n X_i = y\right).$$

Then the joint pmf of (Y, N) is

$$\begin{aligned} p(x,n) &= P(Y = y, N = n) = P\left(\sum_{i=1}^n X_i = y, N = n\right) \\ &= P\left(\sum_{i=1}^n X_i = y\right) P(N = n) = p_n(y)P(N = n). \end{aligned}$$

The conditional pmf is

$$p_{Y|N}(y \mid n) = \frac{p(x,n)}{P(N = n)} = p_n(y).$$

Hence, we have

$$E(Y \mid N = n) = \sum_y y p_n(y) = E\left(\sum_{i=1}^{n} X_i\right) = n\mu_F,$$

where μ_F is the mean of F, so that

$$E(Y) = E\{E(Y \mid N)\} = E(N)\mu_F = \lambda\mu_F.$$

Likewise, we have

$$\operatorname{Var}(Y \mid N = n) = \operatorname{Var}\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} \operatorname{Var}(X_i) = n\sigma_F^2,$$

where σ_F^2 is the variance of F, so that

$$\begin{aligned}\operatorname{Var}(Y) &= \operatorname{Var}\{E(Y \mid N)\} + E\{\operatorname{Var}(Y \mid N)\} \\ &= \operatorname{Var}(N\mu_F) + E(N\sigma_F^2) = \lambda(\mu_F^2 + \sigma_F^2).\end{aligned}$$