Math 408 - Mathematical Statistics

Lecture 6. Expectation, Variance, Covariance, and Correlation

January 30, 2013

Expectation of a Random Variable

The expectation (or mean) of a random variable X is the average value of X. The formal definition is as follows.

Definition

The **expected value**, or **mean**, or **first moment** of X is

$$\mu_X \equiv \mathbb{E}[X] = \begin{cases} \sum_x x f_X(x), & \text{if } X \text{ is discrete} \\ \int x f_X(x) dx, & \text{if } X \text{ is continuous} \end{cases}$$

assuming that the sum (or integral) is well-defined.

Remarks:

- The expectation is a one-number summary of the distribution.
- Think of $\mathbb{E}[X]$ as the average value you would obtain if you computed the numerical average $\frac{1}{n}\sum_{i=1}^{n} X_i$ of a large number of i.i.d. draws $X_1, \ldots, X_n$. The fact that

$$\mathbb{E}[X] \approx \frac{1}{n}\sum_{i=1}^{n} X_i$$

is a theorem called the law of large numbers.

Examples

- Let $X \sim \mathrm{Bernoulli}(p)$. Find $\mathbb{E}[X]$.
  Answer: $\mathbb{E}[X] = p$
- Let $X \sim U(-1, 3)$. Find $\mathbb{E}[X]$.
  Answer: $\mathbb{E}[X] = 1$

Let $Y = r(X)$. How do we compute $\mathbb{E}[Y]$? There are two ways:

- Find $f_Y(y)$ (Lecture 4) and then compute $\mathbb{E}[Y] = \int y f_Y(y) dy$.
- An easier way:

$$\boxed{\mathbb{E}[Y] = \mathbb{E}[r(X)] = \int r(x) f_X(x) dx}$$

Example: Take a stick of unit length and break it at random. Let Y be the length of the longer piece. What is the mean of Y?
Answer: $\mathbb{E}[Y] = \frac{3}{4}$
Functions of several variables are handled in a similar way: if $Z = r(X, Y)$, then

$$\boxed{\mathbb{E}[Z] = \mathbb{E}[r(X, Y)] = \int \int r(x, y) f_{X,Y}(x, y) dx dy}$$

Properties of Expectations

- If $X_1, \ldots, X_n$ are random variables and $a_1, \ldots, a_n$ are constants, then

$$\boxed{\mathbb{E}\left[\sum_{i=1}^{n} a_i X_i\right] = \sum_{i=1}^{n} a_i \mathbb{E}[X_i]}$$

  - Let $X \sim \mathrm{Bin}(n, p)$. Find $\mathbb{E}[X]$.
  - Answer: $\mathbb{E}[X] = np$

- Let $X_1, \ldots, X_n$ be independent random variables. Then,

$$\boxed{\mathbb{E}\left[\prod_{i=1}^{n} X_i\right] = \prod_{i=1}^{n} \mathbb{E}[X_i]}$$

Remark: Note the the summation rule does not require independence but the multiplication rule does.

Variance and Its Properties

The variance measures the "spread" of a distribution.

Definition

Let X be a random variance with mean μ_X.
The **variance** of X, denoted $\mathbb{V}[X]$ or σ_X^2, is defined by

$$\sigma_X^2 \equiv \mathbb{V}[X] = \mathbb{E}[(X-\mu_X)^2] = \begin{cases} \sum_x (x-\mu_X)^2 f_X(x), & \text{if } X \text{ is discrete} \\ \int (x-\mu_X)^2 f_X(x)dx, & \text{if } X \text{ is continuous} \end{cases}$$

The **standard deviation** is $\sigma_X = \sqrt{\mathbb{V}[X]}$

Important Properties of $\mathbb{V}[X]$:

- $\mathbb{V}[X] = \mathbb{E}[X^2] - \mu_X^2$
- If a and b are constants, then $\mathbb{V}[aX+b] = a^2\mathbb{V}[X]$
- If $X_1, \ldots, X_n$ are independent and $a_1, \ldots, a_n$ are constants, then

$$\mathbb{V}\left[\sum_{i=1}^n a_i X_i\right] = \sum_{i=1}^n a_i^2 \mathbb{V}[X_i]$$

Covariance and Correlation

Example: Let $X \sim \mathrm{Bin}(n, p)$. Find $\mathbb{V}[X]$.
Answer: $\mathbb{E}[X] = np(1-p)$

If X and Y are random variables, then the covariance and correlation between X and Y measure how strong the linear relationship is between X and Y.

Definition

Let X and Y be random variables with means μ_X and μ_Y and standard deviations σ_X and σ_Y. Define the **covariance** between X and Y by

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

and the **correlation** by

$$\boxed{\rho(X,Y) = \frac{\mathrm{Cov}(X,Y)}{\sigma_X \sigma_Y}}$$

Properties of Covariance and Correlation

- The covariance satisfies (useful in computations):

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

- The correlation satisfies:

$$-1 \leq \rho(X, Y) \leq 1$$

- If $Y = aX + b$ for some constants a and b, then

$$\rho(X, Y) = \begin{cases} 1, & \text{if } a > 0 \\ -1, & \text{if } a < 0 \end{cases}$$

- If X and Y are independent, then $\mathrm{Cov}(X, Y) = \rho(X, Y) = 0$.
  The converse is not true.
- For random variables $X_1, \ldots, X_n$

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

Expectation and Variance of Important Random Variables

Distribution	Mean	Variance
Point mass at a	a	0
$\mathrm{Bernoulli}(p)$	p	$p(1-p)$
$\mathrm{Bin}(n,p)$	p	$np(1-p)$
$\mathrm{Geom}(p)$	$1/p$	$(1-p)/p^2$
$\mathrm{Poisson}(\lambda)$	λ	λ
$\mathrm{Uniform}(a,b)$	$(a+b)/2$	$(b-a)^2/12$
$\mathcal{N}(\mu,\sigma^2)$	μ	σ^2
$\mathrm{Exp}(\beta)$	β	β^2
$\mathrm{Gamma}(\alpha,\beta)$	$\alpha\beta$	$\alpha\beta^2$
$\mathrm{Beta}(\alpha,\beta)$	$\alpha/(\alpha+\beta)$	$\alpha\beta/((\alpha+\beta)^2(\alpha+\beta+1))$

Summary

- The expected value of X is

$$\mu_X \equiv \mathbb{E}[X] = \begin{cases} \sum_x x f_X(x), & \text{if } X \text{ is discrete} \\ \int x f_X(x) dx, & \text{if } X \text{ is continuous} \end{cases}$$

  - If $Y = r(X)$, then $\mathbb{E}[Y] = \mathbb{E}[r(X)] = \int r(x) f_X(x) dx$
  - If $X_1, \ldots, X_n$ are random variables and $a_1, \ldots, a_n$ are constants, then $\mathbb{E}\left[\sum_{i=1}^n a_i X_i\right] = \sum_{i=1}^n a_i \mathbb{E}[X_i]$
  - If $X_1, \ldots, X_n$ are independent random variables, then $\mathbb{E}\left[\prod_{i=1}^n X_i\right] = \prod_{i=1}^n \mathbb{E}[X_i]$

- The variance of X is

$$\sigma_X^2 \equiv \mathbb{V}[X] = \mathbb{E}[(X - \mu_X)^2]$$

  - $\mathbb{V}[X] = \mathbb{E}[X^2] - \mu_X^2$
  - If a and b are constants, then $\mathbb{V}[aX + b] = a^2 \mathbb{V}[X]$
  - If $X_1, \ldots, X_n$ are independent and $a_1, \ldots, a_n$ are constants, then $\mathbb{V}\left[\sum_{i=1}^n a_i X_i\right] = \sum_{i=1}^n a_i^2 \mathbb{V}[X_i]$

Summary

- Covariance and correlation between X and Y are

$$\boxed{\mathrm{Cov}(X, Y) = \mathbb{E}[(X - \mu_X)(Y - \mu_Y)]} \qquad \boxed{\rho(X, Y) = \frac{\mathrm{Cov}(X, Y)}{\sigma_X \sigma_Y}}$$

  - $\mathrm{Cov}(X, Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]$
  - $-1 \leq \rho(X, Y) \leq 1$
  - If $Y = aX + b$ then $\rho(X, Y) = \begin{cases} 1, & \text{if } a > 0 \\ -1, & \text{if } a < 0 \end{cases}$
  - If X and Y are independent, then $\mathrm{Cov}(X, Y) = \rho(X, Y) = 0$.
  - $\mathbb{V}\left[\sum_{i=1}^n a_i X_i\right] = \sum_{i=1}^n a_i^2 \mathbb{V}[X_i] + 2\sum_{i<j} a_i a_j \mathrm{Cov}(X_i, X_j)$