MATH 3338 Probability

Wenjiang Fu
wfu2@central.uh.edu

Lecture 6 Ch 6. Expected Value and Variance

UNIVERSITY of HOUSTON
DEPARTMENT OF MATHEMATICS

Outline

Chapter 6

Expected values and Variance

Expected value

- **Average value** For a given sequence of finite values $X_1, ..., X_n$, the average value is $\overline{X} = \frac{1}{n}\sum_i X_i$. That's equivalent to say it is the value by giving an equal chance to select each of them.
- **Definition 6.1 (Expected value)** Let X be a numerically valued discrete RV with sample space Ω and distribution function $m(x)$. The expected value $E(X)$ is defined by

$$E(X) = \sum_{x\in\Omega} xm(x),$$

provided this sum converges absolutely. Often, $E(X)$ is called the mean, if the sum converges absolutely. Or we say it does not exist if $\sum_{x\in\Omega} |x|m(x)$ does not converge.

UNIVERSITY of HOUSTON
DEPARTMENT OF MATHEMATICS

- **Example 6.2** Suppose we toss fair coin until a head first comes up. Let X represent the number of tosses to be made. The possible values of X are 1, 2, ... and the distribution of X is defined by $m(i) = 2^{-i}$, i.e. the geometric distribution with prob $p = .5$ We take the expected value of X.

$$E(X) = \sum_{i=1}^{\infty} i \frac{1}{2^i} = \sum_{i=1}^{\infty} [1 + (i-1)] \frac{1}{2^i} = \sum_{i=1}^{\infty} \frac{1}{2^i} + \sum_{i=1}^{\infty} (i-1) \frac{1}{2^i}$$

$$= 1 + \sum_{j=0}^{\infty} j \frac{1}{2^{j+1}} = 1 + \frac{1}{2} \sum_{j=1}^{\infty} j \frac{1}{2^j} = 1 + \frac{1}{2} E(X)$$

Solving for $E(X)$ from the above equation we have $E(X) = 2$.

- **Q**: Why $E(X) = 2$? Can you explain?

Expected value

- **Example 6.3** Suppose we get paid $\$2^{(n-5)}$ each time. Let's find the expected payment.

$$E(Y) = \sum_{i=1}^{\infty} 2^{i-5} \frac{1}{2^i} = \sum_{i=1}^{\infty} 2^{-5} = \infty$$

  Even if we are paid for only even (odd) number of trials, the sum will still diverge.
- **Interpretation of expected value** On the average, the payment or the number of tossing we expect to do.
- **Expectation of a function of RV**
  Theorem 6.1 Discrete RV X has a sample space Ω and distribution function $m(x)$. Function $\phi(x)$ is a real-valued function with domain Ω. Then $\phi(X)$ is a real valued RV. The expected value of $\phi(X)$ is defined by

$$E(\phi(X)) = \sum_{x \in \Omega} \phi(x) m(x)$$

Expected value

- **Sum of Random Variables**
  We may consider the expected value of a sum of RVs or RV multiplied by a constant.
  Theorem 6.2 Let X and XY be rv with finite expected values. Then
  $$E(X+Y) = E(X) + E(Y)$$
  and, if c is any constant, then
  $$E(cX) = cE(X).$$
  Consequently, E is regarded as a linear operator on RVs. For any finite number of RVs $X_1, ..., X_n$, and constants $c_1, ..., c_n$. Then
  $$E(c_1X_1 + ... + c_nX_n) = c_1E(X_1) + ... + c_nE(X_n)$$

UNIVERSITY of HOUSTON
DEPARTMENT OF MATHEMATICS

Expected value

- **Bernoulli Trials**

Theorem 6.3 Let S_n be the sum of n indep Bernoulli trials with the same prob p for a success. Then

$$E(S_n) = np.$$

Proof For each Bernoulli trail, X_i. we have

$$E(X_i) = 1p + 0(1 - p) = p.$$

Then the indep sum of $X_1 ..., X_n$.

$$E(S_n) = E(X_1 + ... + X_n) = E(X_1) + ...E(X_n) = np$$

Notice that S_n follows a Binomial distr $Bin(n, p)$. Hence the mean of a Binomial distr is np.

This is much easier than following the definition of expected value.

Expected value

- **Poisson distribution**
  A Poisson distribution with parameter $\lambda > 0$ has prob

$$P(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}$$

The expected value is

$$E(X) = \sum_{k=0}^{\infty} k \frac{\lambda^k}{k!} e^{-\lambda} = \lambda \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!} e^{-\lambda} = \lambda \sum_{s=0}^{\infty} \frac{\lambda^s}{s!} e^{-\lambda} = \lambda$$

Now consider the multiplication of two RVs.

- **Theorem 6.4** If X and Y are indep rv with finite means, than

$$E(XY) = E(X)E(Y)$$

Expected value

- **Example 6.9**
  A fair coin is tossed twice. Let $X_i = 1$ be the indicator function of a head in the ith toss. We know that they are indep and has $1/2$ as expected value. Then

$$E(X_1X_2) = E(X_1)E(X_2) = 1/4$$

- **Example 6.10** In tossing a coin, let X be the indicator function of head. Define $Y = 1 - X$. Then $XY = 0$ a constant. Thus $E(XY) = 0$. But $E(X)E(Y) = .5 \times .5 = 0.25$. Hence

$$E(XY) \neq E(X)E(Y)$$

- **Example 6.13 - Roulette**
  A roulette wheel has 38 slots numbered 0, 00, 1, 2 ..., 36. 0 and 00 are green, half of the remaining 36 slots are read, and the other half are black. A croupier spins the wheel and throws an ivory ball. Betting: bet \$1 on red, and win \$1 if ball falls in red slot. Or you lose \$1. We calculate the expected value, if you bed \$1 on red.

x	1	-1
m_x	18/38	20 / 38

  Following the definition of expected value,

$$E(X) = 1(18/38) + (-1)(20/38) = -1/19 = -.0526$$

  The interpretation: On the average, players lose \$0.0526 in betting \$1 on the roulette wheel.

UNIVERSITY of HOUSTON
DEPARTMENT OF MATHEMATICS

- **Conditional Expectation**
  Definition 6.2 If F is any event and X is a RV with sample space $\Omega = \{x_1, x_2, ...\}$, then the conditional expectation given F is defined by

$$E(X|F) = \sum_j x_j P(X = x_j|F)$$

  Conditional expectation is used most often in the form provided by the following theorem.

  Theorem 6.5 Let X be a RV with sample space Ω. If $F_1, F_2, ..., F_r$ are events such that $F_i \cap F_j = \emptyset$ for $i \neq j$ and $\Omega = \cup_j F_j$, then

$$E(X) = \sum_j E(X|F_j)P(F_j).$$

  Notice that these events are called mutually exclusive and exhaustive.

Expected value

- **Proof of Theorem 6.5**

$$\sum_j E(X|F_j)P(F_j) = \sum_j \sum_k x_k P(X = x_k|F_j)P(F_j)$$

$$= \sum_j \sum_k x_k P(X = x_k \cap F_j) = \sum_k x_k P(X = x_k) = E(X)$$

UNIVERSITY of HOUSTON
DEPARTMENT OF MATHEMATICS

- **Definition 6.3 - variance**
  Let X be a numerically valued RV with expected value $\mu = E(X)$. Then the variance of X, denoted by $V(X)$, is

$$V(X) = E((X-\mu)^2).$$

  Note that $V(X) = \sum_x (x-\mu)^2 m(x)$, where $m(x)$ is the distribution function of X.
- **Standard deviation** The standard deviation of X, denoted by $D(X)$, or often written as σ, is $D(X) = \sqrt{(X)}$.
  The variance is often denoted by σ^2.
- **Calculation of Variance**
  Theorem 6.6 If X is any RV with $E(X) = \mu$, then

$$V(X) = E(X^2) - \mu^2.$$

  Proof

$$E(X-\mu)^2 = E(X^2 - 2\mu X + \mu^2) = E(X^2) - \mu^2$$

Variance of Discrete RVs

- **Properties of variance**

  Theorem 6.7 If X is any RV and c is any constant, then

$$V(cX) = c^2 V(X)$$

and

$$V(X + c) = V(X)$$

Theorem 6.8 Let X and Y be two independent RVs. Then

$$V(X + Y) = V(X) + V(Y).$$

Theorem 6.9 Let $X_1, X_2, ..., X_n$ be an indep trials process with $E(X_j) = \mu$ and $V(X_j) = \sigma^2$. Let $S_n = X_1 + ... + X_n$ be the sum, and $A_n = S_n/n$ be the average. Then

$$E(S_n) = n\mu, \;\; V(S_n) = n\sigma^2, \;\; \sigma(S_n) = \sigma\sqrt{n}$$

$$E(A_n) = \mu, \;\; V(A_n) = \sigma^2/n, \;\; \sigma(A_n) = \sigma/\sqrt{n}.$$

UNIVERSITY of HOUSTON
DEPARTMENT OF MATHEMATICS

Variance of Discrete RVs

- **Bernoulli trials**

  Let $X_i = 1$ with prob p and 0 with prob $q = 1 - p$. $E(X_i) = p$, $E(X_i^2) = p$. Thus $V(X_i) = p - p^2 = pq$. Further let $S_n = \sum_{i=1}^n X_i$, and $A_n = S_n/n$. Then $E(S_n) = np$. $V(S_n) = npq$, and $D(S_n) = \sqrt{npq}$. $E(A_n) = p$, $V(A_n) = V(S_n)/n^2 = pq/n$, and $D(A_n) = \sqrt{pq/n}$.

  Example 6.19 Let T denote the number of trials until the first success in a Bernoulli trials process. Then T is geometrically distributed. What is the variance of T?

  From previous example, we have

T	1	2	3	...
m_T	p	qp	q^2p	...

  From the geometric dist, $E(T) = 1/p$.

$$E(T^2) = 1p + 4qp + 9q^2p + ... = p(1 + 4q + 9q^2 + ...)$$

Variance of Discrete RVs

- **Example 6.19** (continued)
  In order to figure this out, we need to see
  $$1 + x + x^2 + ... = \frac{1}{1-x}.$$
  Differentiating this, we have
  $$1 + 2x + 3x^2 + ... = \frac{1}{(1-x)^2}$$
  multiplying by x and differentiating again
  $$1 + 4x + 9x^2 + ... = \frac{1+x}{(1-x)^3}.$$
  Thus
  $$E(T^2) = p\frac{1+q}{(1-q)^3} = \frac{1+q}{p^2}$$
  $$V(T) = E(T^2) - (E(T))^2 = \frac{1+q}{p^2} - \frac{1}{p^2} = \frac{q}{p^2}$$

Variance of Discrete RVs

- **Poisson Distribution**

  From previous results, for Poisson distribution, $E(X) = \lambda$. Now we need to calculate $E(X^2)$.

$$E(X^2) = \sum_{x=0}^{\infty} x^2 \frac{\lambda^x}{x!} e^{-\lambda} = \sum_{x=1}^{\infty} x \frac{\lambda^x}{(x-1)!} e^{-\lambda}$$

$$= \sum_{x=1}^{\infty} (x-1) \frac{\lambda^x}{(x-1)!} e^{-\lambda} + \sum_{x=1}^{\infty} \frac{\lambda^x}{(x-1)!} e^{-\lambda}$$

$$= \lambda^2 \sum_{x=2}^{\infty} \frac{\lambda^{x-2}}{(x-2)!} e^{-\lambda} + \lambda \sum_{x=1}^{\infty} \frac{\lambda^{x-1}}{(x-1)!} e^{-\lambda}$$

$$= \lambda^2 + \lambda$$

  Hence $V(x) = E(X^2) - (E(X))^2 = \lambda$.
  i.e., $V(X) = E(X) = \lambda$.

Continuous Random Variables

- **Expected value**
  Definition 6.4 Let X be a real-valued RV with density function $f(x)$. The expected value $\mu = E(X)$ is defined by
  $$\mu = E(X) = \int_{-\infty}^{\infty} xf(x)dx,$$
  provided the integral
  $$\mu = E(X) = \int_{-\infty}^{\infty} |x|f(x)dx,$$
  is finite.
- **Properties of Expected value Theorem 6.10** If X and Y are real-valued RVs and c is any constant. Then
  $$E(X+Y) = E(X) + E(Y), \quad E(cX) = cE(X).$$
  In general,
  $$E(c_1 X_1 + \ldots + c_n X_n) = c_1 E(X_1) + \ldots + c_n E(X_n)$$

- **Expected value**
  Example 6.20 Let X be uniformly distributed on $[0, 1]$. Then The expected value $\mu = E(X)$ is

$$\mu = E(X) = \int_0^1 x dx = 1/2.$$

  It implies that if we choose a large number N of random numbers from $[0, 1]$, and take the average, then we expect that this average should be close to the expected value of 1/2.

- **Example 6.21** Let $Z = (x, y)$ denote a point on the unit disk, and let $X = (x^2 + y^2)^{1/2}$ be the distance from the center. The density function of X can be easily shown to equal $f(x) = 2x$. Why?

$$E(X) = \int_0^1 x f(x) dx = \int_0^1 x x(2x) dx = 2/3.$$

- **Expected value**
  Example 6.22 In an example of a group of people waiting to meet at a hotel, with their arrival time uniformly distributed between 5:00 and 6:00 PM. Let rv Z be the waiting time of the first person who has to wait for the second person to arrive. It has been shown before that $f_Z(z) = 2(1-z)$, for $0 \le z \le 1$. Hence,

$$E(Z) = \int_0^1 zf(z)dz = \int_0^1 2z(1-z)dx = 1/3.$$

- **Expectation of a function of RV Theorem 6.11** If X is a real-valued RV. and if $\phi : \mathbb{R} \to \mathbb{R}$ is a continuous real-valued function with domain $[a, b]$, then

$$E(\phi(X)) = \int_{-\infty}^{+\infty} \phi(x)f_X(x)dx,$$

  provided the integral exists.

- **Expected value**

  Theorem 6.12 Let X and Y be indep real-valued continuous RVs with finite expected values. Then

$$E(XY) = E(X)E(Y).$$

Proof The density of joint distribution and the marginal distributions

$$f_{XY}(x, y) = f_X(x)f_Y(y)$$

Then

$$E(XY) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} xyf_{XY}(x, y)dxdy = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} xyf_X(x)f_Y(y)dxd$$

$$= \int_{-\infty}^{+\infty} xf_X(x)dx \int_{-\infty}^{+\infty} yf_Y(y) = E(X)E(Y)$$

- **Expected value**

  Example 6.24 Let (X, Y) be uniformly distributed on unit square. Let $W = X + Y$. Then Y and W are not indep.

  $$E(y) = 1/2. \;\; E(W) = 1(why?).$$

  $$E(YW) = E(XY+Y^2) = E(X)E(Y)+E(Y^2) = 7/12 \neq E(Y)E(W)$$

- **Variance**

  Definition 6.5 Let X be real-valued RV with density function $f(x)$. The variance $\sigma^2 = Var(X)$ is defined by

  $$\sigma^2 = Var(X) = E((X-\mu)^2).$$

  Theorem 6.13 Let X be real-valued RV with $E(X) = \mu$. Then

  $$\sigma^2 = Var(X) = \int_{-\infty}^{\infty} (x-\mu)^2 f(x) dx.$$

- **Variance**
  Theorem 6.14 Let X be a real-valued RV defined on Ω and c is any constant, then

$$Var(cX) = c^2 Var(X). \;\; Var(X + c) = Var(X).$$

- **Theorems 6.15** If XW is a real-valued RV with $E(X) = \mu$. then

$$Var(X) = E(X^2) - \mu^2.$$

Theorem 6.16 Let X and Y indep RV. Then

$$Var(X + Y) = Var(X) + Var(Y).$$

- **Variance**

  Example 6.26 Let X follow exponential distr with parameter $\lambda > 0$.

$$f(x) = \lambda e^{-\lambda x}$$

$$E(X) = \int_0^\infty x\lambda e^{-\lambda x} dx = 1/\lambda$$

$$Var(X) = E(X^2) - 1/\lambda^2.$$

  Figure out $E(X^2) = 2/\lambda^2$, then $Var(X) = 2/\lambda^2 - 1/\lambda^2 = 1/\lambda^2$.

  Example 6.27 Let Z follow standard normal distribution $N(0, 1)$.

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$$

  $f(x)$ is symmetric about 0, hence,

$$E(X) = \int_{-\infty}^\infty x f_Z(x) dx = 0.$$

- **Variance**
  Example 6.27 Let Z follow standard normal distribution $N(0, 1)$.

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$$

Similarly, we need to work on $E(X^2)$.

$$E(X^2) = \int_{-\infty}^{\infty} x^2 f_Z(x) dx = \sigma^2$$

Then $Var(X) = \sigma^2$. For a general normal distribution RV X, $X = \mu + \sigma Z$, then $E(X) = \mu$, $Var(X) = \sigma^2 Var(Z) = \sigma^2$.

- **Variance**
  Example 6.28 Let X be a continuous RV with Cauchy density function
  $$f(x) = \frac{a}{\pi}\frac{1}{a^2+x^2}$$
  The expectation of X does not exist, because the integral
  $$\frac{a}{\pi}\int_{-\infty}^{\infty}\frac{|x|dx}{a^2+x^2}$$
  diverges. Thus the variance does not exist, either.
- **Independent Trials**
  Corollary 6.1 If $X_1, ..., X_n$ is an indep trials process of real-valued RV with $E(X_i) = \mu$ and $Var(X_i) = \sigma^2$. Let $S_n = X_1 + ... + X_n$, $A_n = S_n/n$. then
  $$E(S_n) = n\mu, \quad E(A_n) = \mu, \quad Var(S_n) = n\sigma^2, \quad Var(An) = \frac{\sigma^2}{n}.$$

- **Variance**
  Corollary 6.1 (continuing) It follows that if we set

$$S_n^* = \frac{S_n - n\mu}{\sqrt{nu\sigma^2}},$$

then

$$E(S_n^*) = 0, \quad Var(S_n^*) = 1$$

In the above, S_n is said to be a standardized version of S_n.

- **Conditional Expectation**
  Recall that we learned conditional expected value (Definition 6.2): For a given event F, random variable X has a conditional expectation
  $$E(X|F) = \sum_j x_j P(X = x_j|F)$$
  Very often, the following is used often with
- **Theorem 6.5** RV X, and a sequence of events $F_1, ..., F_r$ are mutually exclusive and exhaustive $F_i \cap F_j = \emptyset$ for any $i \neq j$, and $\Omega = U_j F_j$. Then
  $$E(X) = \sum_j E(X|F_j)P(F_j)$$

- **Conditional Expectation**
  Let us put this theorem in a slightly different form, and it becomes very useful.
  Conditional Expectation Theorem Assume two random variables X and Y have a joint distribution with $F(x, y)$, can be either discrete or continuous variables. Further we assume the expectation $E(X|Y = y)$ exists for all y in the sample space of Y, and $E(X)$ exists. Then

$$E[E(X|Y)] = E(X)$$

  The proof of the above equality can be done in both discrete and continuous cases.

- **Proof of the Conditional Expectation Theorem** (discrete case)

$$E(X|Y=y) = \sum_i x_i P(X = x_i | Y = y)$$

is a random var on the value of $Y = y$. Taking expectation of this

$$E[E(X|Y=y)] = \sum_y E(X|Y=y)P(Y=y)$$

$$= \sum_j E(X|Y=y_j)P(Y=y_j) = \sum_j \sum_i x_i P(X=x_i|Y=y_j)P(Y=y_j)$$

$$= \sum_i x_i \sum_j P(X=x_i \cap Y=y_j) = \sum_i x_i P(X=x_i) = E(X)$$

- **Example for Conditional Expectation**
  Assume an insurance company receives T number of claims in a week with values X_1, $X_2, ..., X_T$. Assume these claims have a distribution with a mean μ and the number of claims each week has a mean n_0. Calculate the expected value of the total claims. each week.
  The total claims in a week is

$$X_1 + X_2 + ... + X_T = \sum_{i=1}^{T} X_i$$

  Two characteristics:
  1. Each claim amount X_i is a RV with a mean μ.
  2. The number of claims each week T is also a RV with a mean n_0.
  How to calculate the mean total claims? Is the following correct?

$$E[X_1 + X_2 + ... + X_T] = EX_1 + EX_2 + ... + EX_T$$

- **Example for Conditional Expectation**
  The total claims in a week is

$$X_1 + X_2 + ... + X_T = \sum_{i=1}^{T} X_i$$

Solution Use conditional expectation.

$$E\sum_{i=1}^{T} X_i = E[E(\sum_{i=1}^{T} X_i|T)] = E[(\sum_{i=1}^{T} E(X_i|T)] = E[\mu T] = \mu ET = \mu n_0$$

How about conditional variance?

- **Conditional Variance**
  We have defined conditional expectation, we can define conditional variance using conditional expectation.
  Definition of Conditional Variance The conditional variance of X given Y is defined as

$$Var(X|Y) = E(X^2|Y) - [E(X|Y)]^2$$

- **Theorem on Conditional Variance**

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