EE 381 Probability & Statistic with Applications to Computing (Fall 2020)

Lecture 5
Discrete Random Variables (p2)

Duc H. Tran, PhD

Discrete Random Variable's topics

- Discrete probability distributions
- Joint distributions
- **Mathematical expectation**
- **Variance, Standardized random variables**
- Moments and Moment generating function
- Covariance & correlation
- Special probability distributions

Expected Values in Probability

- Consider the experiment: rolling a six-sided dice.
- Repeating it a great many time, what would be the average value of the roll? What is typical deviation from the average?

  use probabilities to compute the values we would expect (meaning that analyzing the experiment beforehand to estimate what will likely happen).
- Two most important expected values in probability:

  Expectation (or Mean)

  Variance

Expectation

- The *expectation* (or *mean)* of a random variable X is denoted by μ_X, and it gives a single value that acts as a representative or average of the values of X.
- For X having the possible values $x_1, x_2, \dots, x_n$, the expectation of X is defined as:

$$E(X) = x_1 P(X = x_1) + \cdots + x_n P(X = x_n) = \sum_{j=1}^{n} x_j \, P\left(X = x_j\right)$$

or equivalently, if $P\left(X = x_j\right) = f(x_j)$:

$$E(X) = x_1 f(x_1) + \cdots + x_n f(x_n) = \sum_{j=1}^{n} x_j \, f\left(x_j\right)$$

- If the probabilities are all equal, we have:

$$E(X) = \frac{x_1 + x_2 + \cdots + x_n}{n}$$

Expectation

Example: Suppose that a game is to be played with a single dice assumed fair. In this game a player wins \$20 if a 2 turns up, \$40 if a 4 turns up; loses \$30 if a 6 turns up; while neither wins or loses if any other faces turn up. Find the expected sum of money (mean) to be won?

- Let X be the random variable giving the amount of money won on any toss. Its probability function is displayed in the below table.

x_j	0	+20	0	+40	0	−30
$f(x_j)$	1/6	1/6	1/6	1/6	1/6	1/6

- The expected sum of money (mean) is:

$$E(X) = (0)\left(\frac{1}{6}\right) + (20)\left(\frac{1}{6}\right) + (0)\left(\frac{1}{6}\right) + (40)\left(\frac{1}{6}\right) + (0)\left(\frac{1}{6}\right) + (-30)\left(\frac{1}{6}\right) = 5$$

Expectation

Example: In a lottery there are 200 prizes of \$5, 20 prizes of \$25, and 5 prizes of \$100. Assuming that 10,000 tickets are to be issued and sold, what is a fair price to pay for a ticket?

- Let X be the random variable denoting the amount of money to be won on a ticket. Its probability function is displayed in the below table.

x (dollars)	5	25	100	0
$P(X = x)$	0.02	0.002	0.0005	0.9775

- For instance, the probability of getting one of the 20 tickets giving a \$25 prize is 20/10000=0.002.
- The expectation of X in dollars is:

$$E(X) = (5)(0.02) + (25)(0.002) + (100)(0.0005) + (0)(0.9775) = 0.2$$

Functions of Random Variables

Let X be a discrete RV with probability function f(x). Then Y=g(X) is also a discrete RV, and the probability function of Y is:

$$h(y) = P(Y = y) = \sum_{\{x|g(x)=y\}} P(X = x) = \sum_{\{x|g(x)=y\}} f(x)$$

If X takes on the values $x_1, x_2, \ldots, x_n$ and Y the values $y_1, y_2, \ldots, y_n$ $(m \leq n)$, then:

$$y_1 h(y_1) + y_2 h(y_2) + \cdots + y_m h(y_m)$$
$$= g(x_1) f(x_1) + g(x_2) f(x_2) + \cdots + g(x_n) f(x_n)$$

Therefore:

$$E[g(X)] = g(x_1)f(x_1) + g(x_2)f(x_2) + \cdots + g(x_n)f(x_n)$$
$$= \sum_{j=1}^{n} g(x_j)f(x_j) = \sum g(x)f(x)$$

Some Theorems on Expectation

Theorem 3-1: If c is any constant, then

$$E(cX) = cE(X)$$

Theorem 3-2: If X and Y are any random variables, then:

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

Theorem 3-3: If X and Y are independent random variables, then:

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

Generalizations can be made easily for all these theorems.

Some Theorems on Expectation

Example: Find the expectation of the sum of points in tossing a pair of fair dice.

Let X and Y be the points showing on the two dice. We have:

$$E(X) = E(Y) = 1\left(\frac{1}{6}\right) + 2\left(\frac{1}{6}\right) + \cdots + 6\left(\frac{1}{6}\right) = \frac{7}{2}$$

Then, by theorem 3-2:

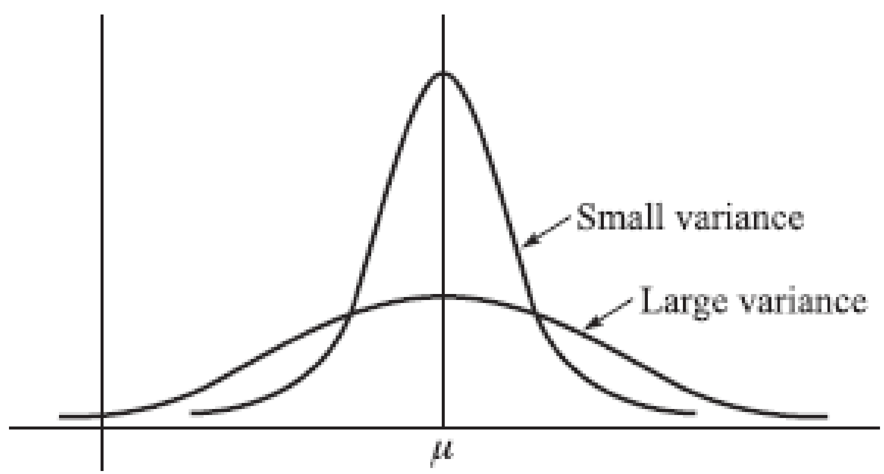
$$E(X+Y) = E(X) + E(Y) = 7$$

The Variance and Standard Deviation

The variance is a measure of the dispersion (how spread out) of the values of the random variable about the mean μ.

If the values tend to be concentrated near the mean, the variance is small.

If the values tend to be distributed far from the mean, the variance is large.



Two continuous distributions having the same mean μ.

The Variance and Standard Deviation

The *variance* is defined by:

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

The variance is a non-negative number. The *standard deviation* is just the square root of the variance.

$$\sigma_x = \sqrt{Var(X)} = \sqrt{E[(X-\mu)^2]}$$

If X is a discrete random variable taking the values $x_1, x_2, \dots, x_n$, and having probability function f(x), then the variance is given by:

$$\sigma_X^2 = E[(X-\mu)^2] = \sum_{j=1}^{n}(x_j-\mu)^2 f(x_j) = \sum (x-\mu)^2 f(x)$$

If the probabilities are all equal, we have:

$$\sigma^2 = [(x_1-\mu)^2 + (x_2-\mu)^2 + \cdots + (x_n-\mu)^2]/n$$

The Variance and Standard Deviation

Example: Rolling a six-sided die. Let X be a random variable denoting the side of the die that is up. X takes on one of the values 1, 2, 3, 4, 5, 6, each with equal probability.

The mean and variance can be computed as:

$$
\begin{aligned}
E[\boldsymbol{X}] &= 1 \cdot \frac{1}{6} + 2 \cdot \frac{1}{6} + 3 \cdot \frac{1}{6} + 4 \cdot \frac{1}{6} + 5 \cdot \frac{1}{6} + 6 \cdot \frac{1}{6} \\
&= 3.5
\end{aligned}
$$

$$
\begin{aligned}
\mathrm{Var}[\boldsymbol{X}] &= (1-3.5)^2 \frac{1}{6} + (2-3.5)^2 \frac{1}{6} + (3-3.5)^2 \frac{1}{6} \\
&\quad + (4-3.5)^2 \frac{1}{6} + (5-3.5)^2 \frac{1}{6} + (6-3.5)^2 \frac{1}{6} \\
&= \frac{2.5^2 + 1.5^2 + 0.5^2 + 0.5^2 + 1.5^2 + 2.5^2}{6} \\
&= \frac{35}{12} = 1.71^2
\end{aligned}
$$

The standard deviation is 1.71.

The Variance and Standard Deviation

Note:

- σ has the same units as random variable X.
- $Var(X)$ has the same units as the square of random variable X. So if X is in meters, then $Var(X)$ is in meters squared.
- Because σ and X have the same units, the standard deviation σ is a natural measure of spread.

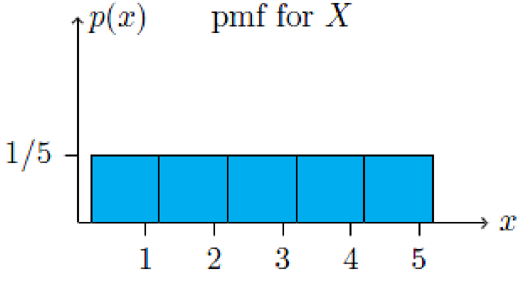
The Variance and Standard Deviation

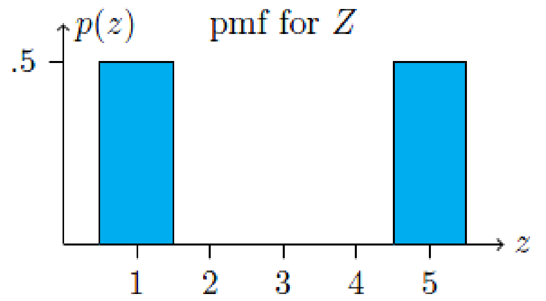
Example: Consider these following random variables

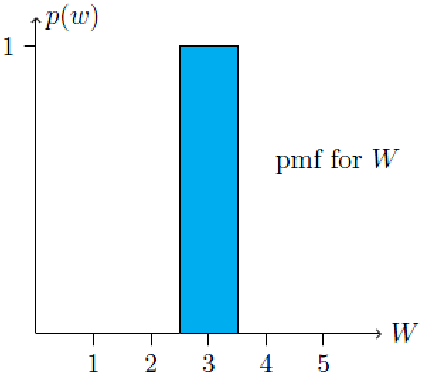
value x	1	2	3	4	5
pmf $p(x)$	$1/5$	$1/5$	$1/5$	$1/5$	$1/5$

value z	1	2	3	4	5
pmf $p(z)$	$5/10$	0	0	0	$5/10$

value w	1	2	3	4	5
pmf $p(w)$	0	0	1	0	0







They all have same mean 3, but the probability is spread out differently.

$$Var(X) = E[(X-\mu)^2] = \frac{4}{5} + \frac{1}{5} + \frac{0}{5} + \frac{1}{5} + \frac{4}{5} = 2$$

$$Var(Z) = E[(Z-\mu)^2] = \frac{20}{10} + \frac{20}{10} = 4$$

$$Var(W) = 0$$

W does not vary, so it has variance 0.

Some Theorem on Variance

Theorem 3-4: $\sigma^2 = E[(X-\mu)^2] = E(X^2) - \mu^2 = E(X^2) - [E(X)]^2$
where $\mu = E(X)$.

Theorem 3-5: If c is any constant

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

Theorem 3-6: The quantity $E[(X-a)^2]$ is a minimum when $a = \mu = E(X)$.

Theorem 3-7: If X and Y are independent random variables,

$$\text{Var}(X+Y) = \text{Var}(X) + \text{Var}(Y) \quad \text{or} \quad \sigma_{X+Y}^2 = \sigma_X^2 + \sigma_Y^2$$

$$\text{Var}(X-Y) = \text{Var}(X) + \text{Var}(Y) \quad \text{or} \quad \sigma_{X-Y}^2 = \sigma_X^2 + \sigma_Y^2$$

Generalization for theorem 3-7 can be easily made.

Some Theorem on Variance

Example: Consider the random variable X, which has mean equals 3:

value x	1	2	3	4	5
pmf $p(x)$	1/5	1/5	1/5	1/5	1/5

$$\mathrm{Var}(X) = E((X-\mu)^2) = \frac{4}{5} + \frac{1}{5} + \frac{0}{5} + \frac{1}{5} + \frac{4}{5} = \boxed{2}.$$

Using theorem 3-4:

$$\sigma^2 = E[(X-\mu)^2] = E(X^2) - \mu^2 = 11 - 9 = 2$$

Formulas of some special sums

$$\sum_{k=1}^{m} k = 1 + 2 + 3 + \cdots + m = \frac{m(m+1)}{2}$$

$$\sum_{k=1}^{m} k^2 = 1^2 + 2^2 + 3^2 + \cdots + m^2 = \frac{(2m+1)(m+1)m}{6}$$

$$\sum_{k=0}^{\infty} r^k = 1 + r + r^2 + \cdots = \frac{1}{1-r} \quad for\ |r| < 1$$

Standardized Random Variables

Let X be the random variable with mean μ and standard deviation σ ($\sigma > 0$). An associated *standardized random variable* is defined by:

$$X^* = \frac{X - \mu}{\sigma}$$

An important property of standardized random variable X^*:

$$E(X^*)=0, \qquad Var(X^*)=1$$

The values of a standardized variable are called *standard scores*, and X is expressed in *standard units*.

Standardized variables are useful for comparing different distributions.

Chebyshev's Inequality

A useful inequality involving probabilities and expected values is Chebyshev's inequality.

Suppose that X is a random variable (discrete or continuous) having mean μ and variance σ^2, which are finite. Then if ϵ is any positive number,

$$P(|X-\mu| \geq \epsilon) \leq \frac{\sigma^2}{\epsilon^2}$$

or with $\epsilon = k\sigma$

$$P(|X-\mu| \geq \epsilon) \leq \frac{1}{k^2}$$

Chebyshev's Inequality

Example: Letting k=2 so that $\epsilon = 2\sigma$ in Chebyshev's inequality, we see that

$$P(|X - \mu| \geq 2\sigma) \leq 0.25$$

or

$$P(|X - \mu| < 2\sigma) \geq 0.75$$

Meaning that the probability of X differing from its mean by more than 2 standard deviation is less than or equal to 0.25. Equivalently, the probability that X will lie within 2 standard deviations of its mean is greater than or equal to 0.75.

Reference

Notes, equations, and figures in the lecture are based on or taken from materials in the course textbook:

“Probability and Statistics”, by Spiegel, Schiller and Srinivasan, ISBN 987-007-179557-9 (McGraw-Hill/Schaun’s)