Basics of discrete distributions

Li Li

Department of Mathematics and Statistics

Random variable

- What is a random variable?
  - A random variable is described informally as a variable whose values depend on outcomes of a random experiment.
  - Typically, a random variable takes numeric values with certain probabilities.
  - Use notations X, Y, Z to denote random variables.
- A random variable is not
  - a number
  - an unknown number

Example

- For the random experiment of tossing a fair coin, define a random variable X as follows

$$X = \begin{cases} 1 & \text{if the coin shows up head} \\ 0 & \text{if the coin shows up tail} \end{cases}$$

- Since the coin is fair, $\underbrace{P(X=1)}_{=P(\text{'head'})} = \underbrace{P(X=0)}_{=P(\text{'tail'})} = 0.5$.

Random variable

- Like random experiment, a random variable also has its sample space.
- The sample space of a random variable could be the same as the sample space of the random experiment.
- For the random variable X defined in the previous slide, the sample space of X is $\{0, 1\}$. We can describe the sample space as

$$X = \{0, 1\}$$

Type of random variables

- Discrete random variables.
- Continuous random variables.

Discrete random variable

- **Discrete random variable**: a discrete random variable is a random variable with a finite or countably infinite range. /sample space
  Notations for discrete random variables are typically capitalized letters, e.g. X, Y, Z.
- **Sample space of a discrete random variable**: $X = \{x_1, x_2, \ldots, x_n\}$.
  - X is the the number of nonconforming connections on a printed circuit board with 10000 connections.
    $X = \{0, 1, 2, \ldots, 10000\}$. $x_1 = 0,\ x_2 = 1,\ x_3 = 2,\ \ldots\ x_{10,000} = 10,000$
  - X is the number of phone calls needed to get connected.
    $X = \{1, 2, \ldots\}$.

Probability mass function

$f(x_1), f(x_2), .. f(x_n)$

- **Probability mass function** *(pmf)*, denoted by $f(x)$, for a discrete random variable X with sample space $\{x_1, x_2, \ldots, x_n\}$ (n could be infinity), is a function such that
  (a) $f(x_i) \stackrel{def}{=} P(X = x_i)$
  (b) $f(x_i) \geq 0$
  (c) $\sum_{i=1}^{n} f(x_i) = 1$

$f(x_1) + f(x_2) + \cdots f(x_n) = 1$
$= P(X \in \{x_1, x_2, \ldots x_n\}) = 1$

- **Probability mass function is often referred as the distribution function of the discrete random variable.**
- For example, if $X = \{1, 2, \ldots, 10\}$, and $f(x) = 1/10$ for $x = 1, 2, \ldots, 10$, then $f(x)$ is a legitimate probability mass function.

$f(1) = \frac{1}{10}, \ f(2) = \frac{1}{10}, \ldots$
$f(10) = \frac{1}{10}$

$P(a < X < b) = P(X \in (a, b)) = \sum_{x_i \in (a,b)} f(x_i)$

$$X = \{\overset{x_1}{1}, \overset{x_2}{2}, \cdots, \overset{x_{10}}{10}\}$$

$$f(x) = \frac{1}{10}, \quad x = 1, 2, \cdots, 10$$

$$P(2 < X < 6) = \sum_{x_i \in (2,6)} f(x_i)$$

$$= f(x_3) + f(x_4) + f(x_5)$$

$$= \frac{3}{10}$$

Example

Choose an adult in the United States at random and ask, "How many days per week do you lift weights?" Call the response X for short. Based on a large sample survey, here is a probability model for the answer you will get:

Days x	0	1	2	3	4	5	6	7
Probability $f(x)$	0.73	0.06	0.06	0.06	0.04	0.02	0.01	0.02

(a) Verify that $f(x)$ is a legitimate probability mass distribution.

(b) Describe the event $X < 4$ in words. What is $P(X < 4)$?

(c) Express the event "lifted weights at least once per week" in terms of X. What is the probability of this event?

Example

$x_1 \quad x_2 \quad x_3 \quad x_8$

$X = \{0, 1, 2, \ldots, 7\}$

$C: \{X \geq 1\}$

a) $f(x_i) = P(X = x_i)$ ✓

$P(X \geq 1) = 1 - P(X = 0)$

$0 \leq f(x_i) \leq 1$ ✓

$\sum_{i=1}^{8} f(x_i) = 1$ ✓

$= 1 - 0.73$

$= 0.27$

b) $X < 4$: the number of days lifting weight is less than 4.

$P(X < 4) = \sum_{x_i < 4} f(x_i) = f(0) + f(1) + f(2) + f(3)$

$= 0.91$

Example

$f(x) = 0, \ x \notin \{1,2,3,4\}$

Suppose a random variable X has probability mass function $f(x) = c(1/2)^x, x = 1, 2, 3, 4$. Determine the following probabilities:

$f(1) = c \cdot \frac{1}{2} \quad f(2) = \frac{c}{4}, \ f(3) = \frac{c}{8}, \ f(4) = \frac{c}{16}$

(a) Find the value of c so that $f(x)$ is a legitimate probability mass function.

(b) $P(X = 1)$. $= f(1) = \frac{16}{15} \cdot \frac{1}{2} = \frac{8}{15}$

(c) $P(X \leq 3)$ $= f(1) + f(2) + f(3) = 1 - f(4) = 1 - \frac{16}{15} \cdot \frac{1}{16} = \frac{14}{15}$

(d) What value of X that is most likely?

a) $\frac{c}{2} + \frac{c}{4} + \frac{c}{8} + \frac{c}{16} = 1 \Rightarrow \frac{8c+4c+2c+c}{16} = \frac{15c}{16} = 1$

$\Rightarrow c = 16/15$

'1' is most likely.

Example

Example of tossing three coins

Consider tossing three identical coins independently and define the random variable X as the number of heads obtained in the three tosses. What is the probability mass function of X?

A Binomial distribution $(3, p)$

Define p as the probability of 'head' for a coin.

$X = \{0, 1, 2, 3\}$.

H T T $\quad p\cdot(1-p)^2$

T H T $\quad p\cdot(1-p)^2$

T T H $\quad p\cdot(1-p)^2$

$$f(0) = P(X=0) = (1-p)\cdot(1-p)\cdot(1-p) = (1-p)^3$$

$$f(1) = 3\cdot p\cdot(1-p)^2 = \binom{3}{1} p\cdot(1-p)^2$$

$$f(2) = \binom{3}{2}\cdot p^2 (1-p)$$

$$f(3) = \binom{3}{3} p^3 = p^3$$

$$F(x) = \begin{cases} 0 & x < 0 \\ (1-p)^3 & 0 \leq x < 1 \\ (1-p)^3 + 3p(1-p)^2 & 1 \leq x < 2 \\ (1-p)^3 + 3p(1-p)^2 + 3p^2(1-p) & 2 \leq x < 3 \\ 1 & 3 \leq x \text{ or } x \geq 3 \end{cases}$$

Cumulative distribution function

F(x) > 0 for any x > min(x_1, ... x_n)

Cumulative distribution function *(cdf)*: the cumulative distribution function of a discrete random variable X, denoted as $F(x)$, is

F(x) = 0 only for x < min(x_1, ... x_n)

$$F(x) = P(X \leq x) = \sum_{x_i \leq x} f(x_i)$$ $-\infty < x < +\infty$

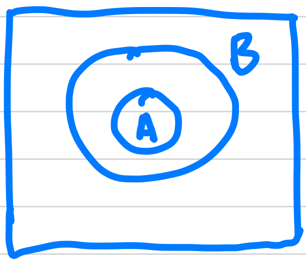
Probability mass function: f(x)

$0 < f(x_i) \leq 1$ for x_i in the sample space

$f(x) = 0$ for x not in the sample space

If $x \uparrow$, $F(x) = \underline{P(X \leq x)} \uparrow$

$\{X \leq x\}$ becomes larger



# of outcomes in B $\geq$

# of outcomes in A

$A \subseteq B$

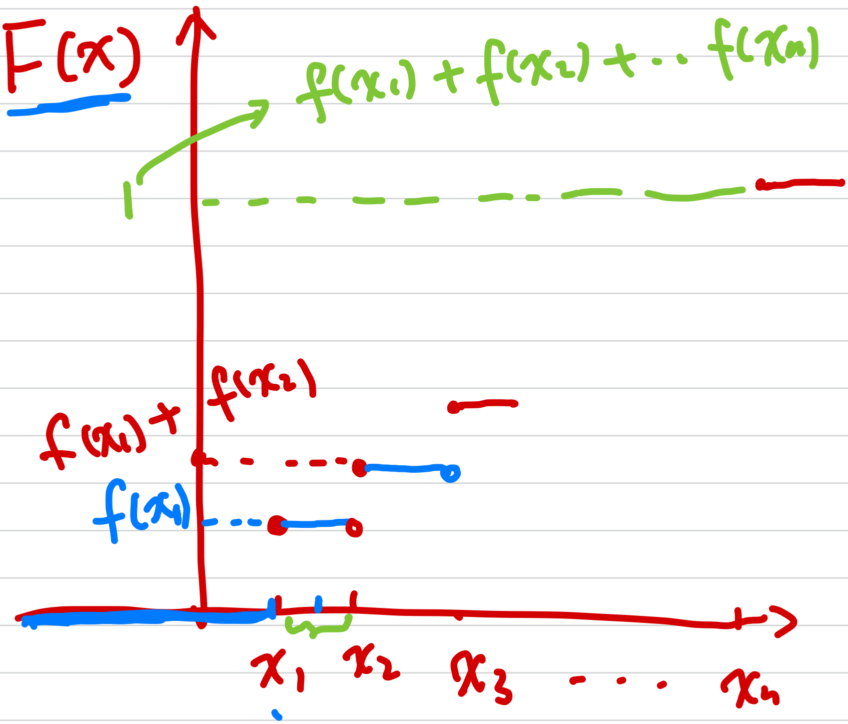
$$P(A) = \sum_{s_i \in A} P(s_i)$$

$$P(B) = \sum_{s_i \in B} P(s_i)$$

$P(A) \leq P(B)$

$$P(X \leq 2) \leq P(X \leq 3)$$

$X = \{x_1, x_2, x_3, \dots, x_n\}$



$$P(X \le x_1) = f(x_1)$$

$$P(X \le x) = \sum_{x_i \le x} f(x_i) \qquad x_1 \le x < x_2$$

$$= f(x_1)$$

$$P(X \le x) = f(x_1) + f(x_2) \qquad x_2 \le x < x_3$$

Example of cdf

If $X = \{1, 2, \ldots, 10\}$, and $f(x) = 1/10$ for $x = 1, 2, \ldots, 10$, then

$$
\begin{aligned}
F(x) &= \sum_{x_i \le x} \frac{1}{n} \\
&= \begin{cases} 0 & x < 1 \\ 1/10 & 1 \le x < 2 \\ 2/10 & 2 \le x < 3 \\ \vdots & \vdots \\ 9/10 & 9 \le x < 10 \\ 1 & x \ge 10 \end{cases}
\end{aligned}
$$

$F(1) = F(1.2) = \frac{1}{10}$

$F(2) = \frac{2}{10}$

$X = \{1, 2, \cdots 10\}$ $f(1) = \frac{1}{10}$

$f(2) = \frac{1}{10}$

Exercise of cdf

Suppose a random variable X has probability mass function $f(x) = (8/7)(1/2)^x, x = 1, 2, 3$. Find the cumulative distribution function of X.

step 1 : $f(x) = \begin{cases} \frac{4}{7} & x=1 \\ \frac{2}{7} & x=2 \\ \frac{1}{7} & x=3 \end{cases}$

step 2 : $F(x) = \begin{cases} 0 & x<1 \\ \frac{4}{7} & 1 \le x < 2 \\ \frac{6}{7} & 2 \le x < 3 \\ 1 & x \ge 3 \end{cases}$

Cumulative distribution function

The cumulative distribution function (CDF) satisfies the following properties:

(a) $0 \leq F(x) \leq 1$

(b) If $x < y$, then $F(x) \leq F(y)$.

Discrete random variables

X_1 : the number of dots for tossing a die

Sample space of X_1 : $X_1 = \{1, 2, 3, 4, 5, 6\}$

X_2 : the number of heads for tossing three coins

$X_2 = \{0, 1, 2, 3\}$

X_3 : the number of cars parked in T for a randomly selected day

$X_3 = \{0, 1, 2, \cdots m\}$ where m is the capacity of the parking lot.

Discrete distributions — probabity mass functions (p.m.f)

$X = \{x_1, x_2, \cdots x_m\}$

p.m.f : $f(x_1), f(x_2), \cdots f(x_m)$

$$f(x_1) = \begin{cases} P(X = x_1) \\ P(X \le x_1) = f(x_1) \\ P(X \cancel{> x_1}) \end{cases}$$

$$f(x_2) = \begin{cases} P(X = x_2) \\ P(X \le x_2) \\ P(X > x_2) \end{cases}$$

$f(x_1) + f(x_2)$

1) $0 \leq f(x_i) \leq 1$

2) $\sum_{i=1}^{m} f(x_i) = 1$; $f(x) = 0$ for $x \notin \{x_1, x_2, \ldots x_m\}$

3) $P(a < X < b) = \sum_{x_i \in (a,b)} f(x_i)$

Bernoulli distribution

x	0	1
$f(x)$	p	$1-p$

$0 < p < 1$

Binomial distribution

$f(x) = \binom{n}{x} p^x (1-p)^{n-x}$ where n, p are constants, $x = 0, 1, 2, \ldots n$.

$f(0) = \binom{n}{0} \cdot p^0 \cdot (1-p)^n$

$f(1) = \binom{n}{1} \cdot p^1 (1-p)^{n-1}$

Cumulative distribution function

$F(x) = P(X \leq x) \quad x \in \mathbb{R}$

when $x \uparrow$ $(-\infty, x] \uparrow$ $F(x)$ is non decreasing

x	0	1
$f(x)$	p	$1-p$

$$F(x) = \begin{cases} 0 & x < 0 \\ p & 0 \leq x < 1 \\ 1 & x > 1 \end{cases}$$

$P(X < -1) = 0 \; ; \; P(X \leq 0) = p$

1) Mean / Expected value

2) Variance

3) Moments

4) General expectation

Distribution mean

Expected value of the discrete distribution or mean of X, denoted by μ or $E(X)$, is defined as

$X = \{x_1 \dots x_n\}$

$\rightarrow$ Expected

$$x_1 f(x_1) + x_2 f(x_2) + \cdots + x_n f(x_n).$$

- Interpretation of mean: in very many repetitions of the experiment setting, mean is the average value of X.
- For example, the expected value in rolling a six-sided die is 3.5, because the average of all the numbers that come up in an extremely large number of rolls is close to 3.5.
- Less roughly, the law of large numbers states that the arithmetic mean of the values almost surely converges to the expected value as the number of repetitions approaches infinity.

For a fair six-faced die

$X = \{1, 2, 3, 4, 5, 6)$

$f(x_i) = \frac{1}{6}$

$$E(X) = 1 \cdot \frac{1}{6} + 2 \cdot \frac{1}{6} + 3 \cdot \frac{1}{6} + 4 \cdot \frac{1}{6} + 5 \cdot \frac{1}{6} + 6 \cdot \frac{1}{6}$$

$$= \frac{(1 + 2 + \cdots 6)}{6} = \frac{21}{6} = \frac{7}{2} = 3.5$$

Distribution mean

To compute the expected value for a discrete distribution, we take the weighted average of the sample space for the distribution. For example, if $X = \{1, 2, \ldots, 10\}$, and $f(x) = 1/10$ for $x = 1, 2, \ldots, 10$, then

$$E(X) = 1 * \frac{1}{10} + 2 * \frac{1}{10} + \cdots + 10 * \frac{1}{10} = \frac{1}{10}(1 + 2 + \cdots + 10) = \frac{55}{10} = 5.5.$$

Example

Your company employs 1,000 sales representatives. You are considering whether or not to hire a sales assistant for each sales representative. You are given the following data: Additional total company revenue from 1,000 sales assistants

X Revenue (million dollars)	f(x) Probability
0	20%
25	30%
100	30%
200	20%

$E(X) = 0 \cdot 0.2 + 25 \cdot 0.3 + 100 \cdot 0.3 + 200 \cdot 0.2 = 77.5$

In addition, cost of a sales assistant is 30, 000 and one-time fixed hiring costs is 3 million.

(a) Calculate the expected revenue from a 1000 sales assistants.

$Cost = 0.03 \cdot 1000 + 3 = 33$

(b) Explain how the expected revenue can help make the decision of whether to hire 1000 sales assistants.

Example

Where do we use mean in Statistics?

Examples:

(a) Determine the average performance of SAT takers. The average score is the mean of the distribution for all SAT scores.

(b) Comparing the average effects of two treatments, for example, the average amount of decreases in cholesterol of two drugs. The average treatment effect of a drug is the mean of the distribution for cholesterol decrease across the population.

(c) Determine the averaged effect of a factor on the response. For example, the average amount of change in annual household income when tax rate decreases by 1%. The averaged effect of the factor is the mean of the distribution for changes in annual household income when tax rate decreases by 1%.

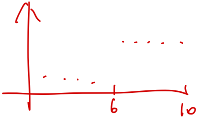
Distribution variance

Variance of the discrete distribution or variance of X, denoted by σ^2 or $E(X-\mu)^2$, is defined as

$$(x_1-\mu)^2 f(x_1) + (x_2-\mu)^2 f(x_2) + \cdots + (x_n-\mu)^2 f(x_n).$$

- Interpretation of variance: variance is the expectation of the squared deviation of a random variable from its mean.
- Informally, it measures how far a set of (random) numbers are spread out from their average value.

Distribution variance



To compute the variance for a discrete distribution, we take the weighted average of squared deviation of a random variable from its mean.

For example, if $X = \{1, 2, \ldots, 10\}$, and $f(x) = 1/10$ for $x = 1, 2, \ldots, 10$, then

$\mu = 5.5$

$$Var(X) = (1-5.5)^2 * \frac{1}{10} + (2-5.5)^2 * \frac{1}{10} + \cdots + (10-5.5)^2 * \frac{1}{10} = 8.25$$

2) $X = \{1, 2, \ldots 10\}$. $f(x) = \begin{cases} \frac{1}{20} & x = 1, 2, \ldots 5 \\ \frac{3}{20}, & x = 6, 7, \ldots 10 \end{cases}$

$\mu = \frac{1+2+3+4+5}{20} + \frac{3}{20} \cdot (6+7+8+9+10) = 6.75$

$\sigma^2 = (1-6.75)^2 \cdot \frac{1}{20} + \ldots (5-6.75)^2 \cdot \frac{1}{20} + (6-6.75)^2 \cdot \frac{3}{20} + \ldots (10-6.75)^2 \cdot \frac{3}{20}$

Equivalent representation of the formula for variance

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

$\sigma^2 = (x_1-\mu)^2 \cdot f(x_1) + (x_2-\mu)^2 \cdot f(x_2) + \cdots$ $(x_n-\mu)^2 \cdot f(x_n)$

$$\sigma^2 = E(X^2) - \mu^2 = x_1^2 f(x_1) + x_2^2 f(x_2) + \cdots + x_n^2 f(x_n) - \mu^2.$$

For example, if $X = \{1, 2, \ldots, 10\}$, and $f(x) = 1/10$ for $x = 1, 2, \ldots, 10$, then

$$Var(X) = 1^2 \frac{1}{10} + 2^2 \frac{1}{10} + \ldots 10^2 \frac{1}{10} - 5.5^2 = 8.25.$$

Where do we use variance in Statistics?

Examples:

(a) Determine the risk of a portfolio of stocks. We characterize the risk by the variance of the distribution for the gains of the portfolio.

(b) Comparing the average effects of two treatments needs to rescale the treatment effect of each subject using the variance of the distribution for the treatment effect.

(c) Analysis of variance (ANOVA) is a collection of statistical models and their associated estimation procedures used to analyze the differences among group means in a sample. In the ANOVA setting, the observed variance in a particular variable is partitioned into components attributable to different sources of variation.

Exercise

Suppose a random variable X has probability mass function $f(x) = (8/7)(1/2)^x, x = 1, 2, 3$. Find the mean and variance of the distribution.

$E(X) = \frac{1}{3}(1+2+3) = 2$ X

$Var(X) = (1-2)^2 \cdot \frac{1}{3} + (2-2)^2 \cdot \frac{1}{3} + (3-2)^2 \cdot \frac{1}{3} = \frac{2}{3}$ X

$E(X) = \frac{4}{7} \cdot 1 + \frac{2}{7} \cdot 2 + \frac{1}{7} \cdot 3 = \frac{11}{7}$

$Var(X) = (1-\frac{11}{7})^2 \cdot \frac{4}{7} + (2-\frac{11}{7})^2 \cdot \frac{2}{7} + (3-\frac{11}{7})^2 \cdot \frac{1}{7}$

$= 1^2 \cdot \frac{4}{7} + 2^2 \cdot \frac{2}{7} + 3^2 \cdot \frac{1}{7} - (\frac{11}{7})^2$

$= 0.5306 = \frac{19}{7} - \frac{121}{49} = \frac{26}{49}$

Standard deviation : σ

$$\sigma = \sqrt{\sigma^2} \rightarrow \text{variance}$$

Weight average distance between possible x values and the distribution mean.

Moments

$E(X^2)$

- **Moments of a discrete distribution or of** X, denoted by $E(X^k), k = 1, 2, \ldots$ are called the first moment, second moment, etc. Moments are calculated as

$$x_1^k f(x_1) + x_2^k f(x_2) + \cdots + x_n^k f(x_n).$$

First moment: $x_1 f(x_1) + x_2 f(x_2) + \ldots x_n f(x_n) = \mu$

- They are also called expected value of X^k.
- The first moment is the mean of the distribution and the second moment minus the first moment square is the variance of the distribution.

$E(X^2) = x_1^2 \cdot f(x_1) + x_2^2 \cdot f(x_2) + \cdots x_n^2 \cdot f(x_n)$

- For example, if $X = \{1, 2, \ldots, 10\}$, and $f(x) = 1/10$ for $x = 1, 2, \ldots, 10$, then

$$E(X^3) = 1^3 * \frac{1}{10} + 2^3 * \frac{1}{10} + \ldots 10^3 * \frac{1}{10} = 302.5.$$

Exercise

Suppose a random variable X has probability mass function $f(x) = (8/7)(1/2)^x, x = 1, 2, 3$. Find the third moment of the distribution.

General expectations

- For a discrete distribution or a discrete random variable X, expectation of a function of X is defined as

$$E(g(X)) = g(x_1)f(x_1) + g(x_2)f(x_2) + \cdots + g(x_n)f(x_n).$$

- For example, if $X = \{1, 2, \ldots, 10\}$, and $f(x) = 1/10$ for $x = 1, 2, \ldots, 10$,
$E(\sqrt{X}) = 1 * \frac{1}{10} + \sqrt{2} * \frac{1}{10} + \cdots + \sqrt{10} * \frac{1}{10} = 2.25.$

Exercise

Suppose a random variable X has probability mass function $f(x) = (8/7)(1/2)^x, x = 1, 2, 3$. Find $E(2X+1)$.

$g(X) = 2X+1$

X	$1 \; (x_1)$	$2 \; (x_2)$	$3 \; (x_3)$
$f(x)$	$\frac{4}{7}$	$\frac{2}{7}$	$\frac{1}{7}$

$E(X) = \frac{11}{7}$

$E(2X+1) = 2 \cdot E(X) + 1$
$= \frac{29}{7}$

$E(2X+1)$
$= \sum_{i=1}^{n} (2x_i+1) f(x_i)$
$= (2\cdot 1+1)\cdot\frac{4}{7} + 5\cdot\frac{2}{7} + 7\cdot\frac{1}{7}$
$= \frac{29}{7}$

Moment generating function

- The moment-generating function of a real-valued random variable is an alternative specification of its probability distribution. Thus, it provides the basis of an alternative route to analytical results compared with working directly with probability density functions or cumulative distribution functions.
- However, not all random variables have moment-generating functions.
- **Moment generating function of a discrete distribution or of X:**

$$M(t) = E(e^{tX}) = e^{tx_1} f(x_1) + e^{tx_2} f(x_2) + \cdots + e^{tx_n} f(x_n)$$

Moment generating function

For example, if $X = \{1, 2, \ldots, 10\}$, and $f(x) = 1/10$ for $x = 1, 2, \ldots, 10$, then

$$M(t) = e^t * \frac{1}{10} + e^{2t} * \frac{1}{10} + \cdots + e^{10t} * \frac{1}{10} = \frac{1}{10}(e^t + e^{2t} + \cdots + e^{10t}).$$

Moment generating function

For example, if $X = \{1, 2, \ldots, 10\}$, and $f(x) = 1/10$ for $x = 1, 2, \ldots, 10$, then

$$M(t) = e^t * \frac{1}{10} + e^{2t} * \frac{1}{10} + \cdots + e^{10t} * \frac{1}{10} = \frac{1}{10}(e^t + e^{2t} + \cdots + e^{10t}).$$

Finding moments using the moment generating function

If a moment-generating function exists for a random variable X, then:

(a) The mean of X can be found by evaluating the first derivative of the moment-generating function at $t = 0$. That is: $\mu = M'(0)$

(b) The variance of X can be found by evaluating the first and second derivatives of the moment-generating function at $t = 0$. That is: $\sigma^2 = M''(0) - (M'(0))^2$.

(c) It is not always more convenient to use moment generating function to obtain moments.

Example

If $X = \{1, 2, \ldots, 10\}$, and $f(x) = 1/10$ for $x = 1, 2, \ldots, 10$, then

$$\begin{aligned} M'(t) &= \frac{1}{10}(e^t + 2e^{2t} + \cdots + 10e^{10t}) \\ M''(t) &= \frac{1}{10}(e^t + 2^2e^{2t} + \cdots + 10^2e^{10t}) \end{aligned}$$

$$\begin{aligned} M'(0) &= \frac{1}{10}(1 + 2 + \cdots + 10) = 5.5 \\ M''(0) &= \frac{1}{10}(1 + 2^2 + \cdots + 10^2) = 38.5 \end{aligned}$$

Hence $\mu = 5.5$ and $\sigma^2 = 38.5 - 5.5^2 = 8.25$.

Exercise

Suppose a random variable X has probability mass function $f(x) = (8/7)(1/2)^x, x = 1, 2, 3$. Find $M(t)$ for the random variable X and use it to derive $E(X)$ and $Var(X)$.

$E(X) = \frac{11}{7}$; $Var(X) = \frac{26}{49}$

$M(t) = e^t \cdot \frac{4}{7} + e^{2t} \cdot \frac{2}{7} + e^{3t} \cdot \frac{1}{7}$

$M'(t) = e^t \cdot \frac{4}{7} + e^{2t} \cdot \frac{4}{7} + e^{3t} \cdot \frac{3}{7}$

$M'(0) = \frac{4}{7} + \frac{4}{7} + \frac{3}{7} = \frac{11}{7}$

$M''(t) = e^t \cdot \frac{4}{7} + e^{2t} \cdot \frac{8}{7} + e^{3t} \cdot \frac{9}{7}$

$M''(0) = \frac{21}{7}$; $Var(X) = \frac{21}{7} - \left(\frac{4}{7}\right)^2 = \frac{26}{49}$

Properties of expectation

Let X be a random variable and let a, b, and c be constants. Providing the following expectations exists,

(a) $E(aX + c) = aE(X) + c$

(b) $E(aX^2 + bX + c) = aE(X^2) + bE(X) + c$.

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

$$= E(X^2) - 2\mu \underbrace{E(X)}_{\mu} + \mu^2$$

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

Example

Suppose a random variable X has probability mass function $f(x) = (8/7)(1/2)^x, x = 1, 2, 3$. Find $E(2X + 1)$ using property (a) and $E(X^2 - 2X + 1)$ using property (b).

$E(X) = \frac{11}{7}$; $E(X^2) = \frac{21}{7}$

$E(2X+1) = 2 \cdot E(X) + 1 = 2 \cdot \frac{11}{7} + 1 = \frac{29}{7}$

$E(X^2 - 2X + 1) = E(X^2) - 2E(X) + 1$

$= \frac{21}{7} - \frac{22}{7} + 1 = \frac{6}{7}$

More properties

Let $g_1(X), g_2(X), \ldots, g_k(X)$ be functions of random variable X, then

$$E(g_1(X)+g_2(X)+\cdots+g_k(X)) = E(g_1(X))+E(g_2(X))+\cdots+E(g_k(X))$$

where k is an integer.

$$E(\underbrace{aX^2}_{g_1}+\underbrace{bx}_{g_2}+\underbrace{c}_{g_3}) = E(aX^2)+E(bx)+E(c)$$

$$= a\,E(X^2)+bE(X)+c$$

Exercise

Suppose a random variable X has probability mass function $f(x) = (8/7)(1/2)^x, x = 1, 2, 3$. Find $E(\sqrt{X} + X^2 + 1)$.