Math 3501 - Probability and Statistics I

2.3 - Special Mathematical Expectations

Moment generating function

Definition

Let X be a random variable of the discrete type with pmf $f(x)$ and space S.

If there is a positive number h such that

$$E\left(e^{tX}\right) = \sum_{x \in S} e^{tx} f(x)$$ function of t

exists and is finite for $-h < t < h$, then the function defined by

$$\boxed{M(t) = E\left(e^{tX}\right), \quad t \in (-h, h), \; h > 0}$$

is called the *moment-generating function of* X (or of the distribution of X).

We often use the abbreviation mgf.

For $t=0$ $E[e^{tX}] = E[1] = 1$ is defined

Question is whether $E[e^{tX}]$ is defined for $t \neq 0$

Notes:

$M(t) = \sum_{x \in S} e^{tx} \cdot f(x)$

1) If $S = \{b_1, b_2, b_3, \ldots\}$, then the mgf is given by the expansion

$$M(t) = e^{tb_1} f(b_1) + e^{tb_2} f(b_2) + e^{tb_3} f(b_3) + \cdots ,$$

where the coefficient of e^{tb_i} is the probability

$$f(b_i) = P(X = b_i) .$$

2) The moment-generating function of a discrete random variable uniquely determines the distribution of that random variable:

- if the mgf exists, there is one and only one distribution of probability associated with that mgf.

$\left.\begin{array}{l} \text{pmf} \\ \text{cdf} \\ \text{mgf} \end{array}\right\} \rightarrow$ uniquely specify the distribution!

Example

Suppose X has the mgf

$$M(t) = e^t\left(\frac{3}{6}\right) + e^{2t}\left(\frac{2}{6}\right) + e^{3t}\left(\frac{1}{6}\right), \quad -\infty < t < \infty$$

Indicate the support of X and the associated probabilities.

Recall that if $S = \{b_1, b_2, b_3, \dots\}$ then

$$M(t) = E[e^{tX}] = \sum_{x \in S} e^{tx} \cdot f(x) = e^{tb_1} \cdot f(b_1) + e^{tb_2} f(b_2) + e^{tb_3} \cdot f(b_3) + \dots$$

In our example, we have only three exponential terms $\Rightarrow$ Space of X : S has only 3 elements

$S = \{1, 2, 3\}$ is the support of X

numbers multiplying t on the exponentials

and the corresponding probabilities are:

$f(1) = 3/6$, $f(2) = \frac{2}{6}$, $f(3) = 1/6$

Very useful property: → the reason why we call $M(t) = E[e^{tX}]$ the moment generating function :

If the mgf exists, then for each positive integer r, we have

$$\boxed{M^{(r)}(0) = \sum_{x \in S} x^r f(x) = E(X^r)\ .}$$

From the definition, we have $M(t) = E[e^{tX}] = \sum_{x \in S} e^{tx} \cdot f(x)$ whenever mgf exists

Taking the derivative with respect to t :

$$M'(t) = \sum_{x \in S} x\, e^{tx} \cdot f(x) \overset{t=0}{\Longrightarrow} M'(0) = \sum_{x \in S} x \cdot f(x) = E[X]$$

because $e^0 = 1$

$$M''(t) = \sum_{x \in S} x^2 e^{tx} f(x) \Rightarrow M''(0) = \sum_{x \in S} x^2 f(x) = E[X^2]$$

$\vdots \qquad \vdots$

For any $r \geq 1$; $$M^{(r)}(t) = \sum_{x \in S} x^r e^{tx} f(x) \Rightarrow M^{(r)}(0) = \sum_{x \in S} x^r f(x) = E[X^r]$$

we can find moments about the origin of any order by differentiating $M(t)$ w.r.t t (the appropriate # of times) and setting $t=0$

Example

Suppose X has a geometric distribution, that is, the pmf of X is given by

$$f(x) = q^{x-1}p\,, \quad x = \underbrace{1, 2, 3, \dots}_{\text{support is infinite set}}\,,$$

where $q = 1 - p$ $\Rightarrow$ $1 - q = p$

Find the mfg of X and use it to determine the mean and variance of X.

The mgf of X is

$$M(t) = E[e^{tX}] = \sum_{x \in S} e^{tx} \cdot f(x) = \sum_{x=1}^{\infty} e^{tx}\, \overbrace{q^{x-1} \cdot p}^{f(x)} = ???$$

$$= p \cdot q^{-1} \sum_{x=1}^{\infty} \underbrace{e^{tx} \cdot q^x}_{(e^t)^x \cdot q^x} = \frac{p}{q} \underbrace{\sum_{x=1}^{\infty} (e^t \cdot q)^x}_{} = \frac{p}{\cancel{q}} \cdot \frac{\cancel{q} e^t}{1 - q e^t}$$

provided $e^t \cdot q < 1$

geometric series with first term $e^t \cdot q$ and ratio $e^t \cdot q$

$\Rightarrow$ convergent provided $\overbrace{|e^t \cdot q|}^{e^t \cdot q} < 1$

$$= \frac{p e^t}{1 - q e^t}$$

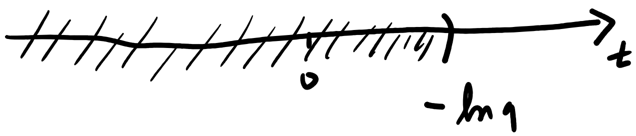
Conclusion: m.g.f of X is

$$\underbrace{M(t) = \frac{pe^t}{1-qe^t}}_{\text{defined for all } t < \underbrace{-\ln q}_{\text{positive because } q>0 \text{ and so } \ln q<0}} , \quad \text{for } t \text{ s.t. } \underbrace{e^t \cdot q < 1}$$

$$e^t < \frac{1}{q}$$

$$t < \ln \frac{1}{q}$$

$$t < -\ln q$$



We can evaluate $\underbrace{E[X]}_{\mu}$ and $E[X^2]$ (so that we can find $Var(X) = E[X^2] - (E[X])^2$)

Recall that

$$E[X] = M'(0) \qquad \text{and} \qquad E[X^2] = M''(0)$$

Then, from $M(t) = \dfrac{pe^t}{1-qe^t}$, we get

the term $pq\,e^{2t}$ cancels with $-pq\,e^{2t}$

$$M'(t) = \frac{pe^t(1-qe^t) - pe^t\cdot(-qe^t)}{(1-qe^t)^2} = \frac{pe^t}{(1-qe^t)^2}$$

quotient rule!

$$\Rightarrow E[X] = M'(0) = \frac{p}{(1-q)^2} = \frac{p}{p^2} = \frac{1}{p}$$

$$M''(t) = \frac{pe^t(1-qe^t)^2 - pe^t\cdot 2(-qe^t)(1-qe^t)}{(1-qe^t)^4}$$

$$= \frac{(1-qe^t)\left[pe^t(1-qe^t) + 2pe^t qe^t\right]}{(1-qe^t)^4} = \frac{pe^t \overbrace{-pqe^{2t} + 2pqe^{2t}}^{pqe^{2t}}}{(1-qe^t)^3}$$

$$= \frac{pe^t(1+qe^t)}{(1-qe^t)^3}$$

$$\Rightarrow E[X^2] = \frac{p(1+q)}{(1-q)^3} = \frac{p(1+q)}{p^3} = \frac{1+q}{p^2}$$

The alternative would be to evaluate the sum $E[x^2] = \sum_{x=1}^{\infty} x^2 \cdot q^{x-1} \cdot p$

Conclusion:

mean $\mu = E[X] = \frac{1}{p}$

Variance $Var(X) = E[X^2] - (E[X])^2 = \frac{1+q}{p^2} - \left(\frac{1}{p}\right)^2 = \frac{q}{p^2}$

Math 3501 - Probability and Statistics I

2.4 - The binomial distribution

Bernoulli experiment

A *Bernoulli experiment* or *Bernoulli trial* is a random experiment, the outcome of which can be classified in one of two mutually exclusive and exhaustive ways, usually referred to as *success* or *failure*.

- heads or tails
- female or male
- life or death
- nondefective or defective

Bernoulli trials

A sequence of *independent Bernoulli trials* occurs when a Bernoulli experiment is performed several independent times and the probability of success, usually denoted as p, remains the same from trial to trial.

Very often:

- p denotes the probability of success on each trial
- $q = 1 - p$ denote the probability of failure

Example

Suppose that the probability of germination of a beet seed is 0.8 and the germination of a seed is called a success.

If we plant 10 seeds and can assume that the germination of one seed is independent of the germination of another seed, this would correspond to 10 Bernoulli trials with $p = 0.8$.

Bernoulli distribution

Let X be a random variable associated with a Bernoulli trial, defined as:

$$X(\text{success}) = 1 \qquad \text{and} \qquad X(\text{failure}) = 0 \; .$$

pmf of X is

$$f(x) = \begin{cases} 1-p & \text{if } x = 0 \\ p & \text{if } x = 1 \end{cases}$$

The pmf of X is

$$\boxed{f(x) = p^x(1-p)^{1-x} \; , \quad x = 0, 1}$$

and we say that X has a *Bernoulli distribution* (with parameter p).

Bernoulli distribution mean and variance

Let X be a random variable with a Bernoulli distribution with parameter p. Then:

- $\mu = E[X] = p$ ✓
- $\sigma^2 = \text{Var}(X) = pq$ ✓

$$\mu = E[X] = \sum_{x=0,1} x \cdot f(x) = 0 \cdot f(0) + 1 \cdot f(1) = 0 \cdot (1-p) + 1 \cdot p = p$$

$$\sigma^2 = \text{Var}(X) = E\left[(X - E[X])^2\right] = E\left[(X-p)^2\right] = \sum_{x=0,1} (x-p)^2 \cdot f(x)$$

$$= (0-p)^2 \cdot f(0) + (1-p)^2 \cdot f(1) = p^2 \cdot (1-p) + (1-p)^2 \cdot p =$$

$$= (1-p)\, p \underbrace{\left[p + (1-p) \right]}_{=1} = p(1-p) = p \cdot q$$

we could also have computed the variance using $\text{Var}(X) = E[X^2] - (E[X])^2$ but it wouldn't make it a lot simpler

Random sample from a Bernoulli distribution

In a sequence of n Bernoulli trials, denote by X_i the Bernoulli random variable associated with the ith trial.

An observed sequence of n independent Bernoulli trials is an n-tuple of 0s and 1s.

This sequence is often called a *random sample of size n from a Bernoulli distribution*.

3 1s and 2 0s

$$\frac{5!}{3!\,2!}$$

Example

1 1 1 0 0

1 0 0 1 1

If five beet seeds are planted in a row, a possible observed sequence would be

$n=5$

$$(1, 0, 1, 0, 1)$$

$\overbrace{\binom{3}{2}} \cdot \underbrace{p^3 \cdot (1-p)^2}$ → prob of observing 3 successes

in which the first, third, and fifth seeds germinated and the other two did not.

If the probability of germination is $p = 0.8$, the probability of this outcome is, assuming independence,

successes observed out of 5

of failures observed out of 5

$$(0.8)(0.2)(0.8)(0.2)(0.8) = (0.8)^3(0.2)^2 .$$

In a sequence of Bernoulli trials, we are often interested in the total number of successes but not the actual order of their occurrences.

Let X equal the number of observed successes in n Bernoulli trials.

The possible values of X are $0, 1, 2, \ldots, n$.

- If x successes occur, where $x = 0, 1, 2, \ldots, n$, then $n - x$ failures occur.

The number of ways of selecting x positions for the x successes in the n trials is

$$\binom{n}{x} = \frac{n!}{x!(n-x)!}$$

# of distinguishable permutations of x 1s and $n-x$ 0s

$$\overbrace{1 \quad 0 \quad 1 \quad 1 \quad 0 \ \ldots\ 1}^{n \text{ repetitions}}$$

x 1s (successes)

$n-x$ 0s (failures)

Since:

- trials are independent
- probabilities of success and failure are, respectively, p and $q = 1 - p$

the probability of each of these ways is

successes

$$p^x (1-p)^{n-x} .$$

failures

Binomial distribution

Let X equal the number of observed successes in n Bernoulli trials.

The pmf of X is $P(X = x)$ ← probability of observing x successes in n Bernoulli trial

total # of sequences of 1s and 0s of length n with x 1s and $n-x$ 0s

$$\rightarrow \quad f(x) = \binom{n}{x} p^x (1-p)^{n-x}, \quad x = 0, 1, 2, \ldots, n$$

These probabilities are called *binomial probabilities*, and the random variable X is said to have a *binomial distribution*.

Notation and Terminology:

repetitions

probability of success

- A binomial distribution is denoted by $b(n, p)$ and we sometimes write

$$\boxed{X \sim b(n, p)} \quad \leftarrow \text{NOTATION}$$

  to denote that the distribution of X is $b(n, p)$
- The constants n and p are called the parameters of the binomial distribution

$$\sum_{x=0}^{n} \binom{n}{x} = 2^n$$

Recall the binomial theorem: for any integer $n \geq 0$, it holds

$$(a+b)^n = \sum_{x=0}^{\infty} \left(\begin{array}{c} n \\ x \end{array} \right) a^x b^{n-x} .$$

As a consequence, the pmf of the binomial distribution satisfies

$$\underbrace{\sum_{x=0}^{n} f(x)}_{\text{pmf of } b(n,p)} = \underbrace{\sum_{x=0}^{n} \left(\begin{array}{c} n \\ x \end{array} \right) p^x (1-p)^{n-x}}_{a=p,\ b=1-p} \underset{\text{binomial thm}}{=} (\overset{a}{p} + \overset{b}{(1-p)})^n = 1$$

$p + (1-p) = 1$

$\Rightarrow \quad f(x) = \binom{n}{x} p^x (1-p)^{n-x}, \quad x = 0, 1, \ldots, n$ is a pmf

Summary

A binomial experiment satisfies the following properties:

1. A Bernoulli (success-failure) experiment is performed n times, where n is a (non-random) constant.

2. The Bernoulli trials are independent.

3. The probability of success on each trial is a constant p; the probability of failure is $q = 1 - p$.

4. The random variable X equals the number of successes in the n trials.

$X \sim bi(n, p)$ means X = # successes in n independent Bernoulli trials, each with prob. of success p

Example

An instant lottery has 30% of winning tickets. Assume independence among winning and losing tickets.

Knowing that 8 tickets were purchased, find the probability that

a) - exactly 2 are winning tickets;
b) - at most 3 are winning tickets;
c) - more than 2 are winning tickets.

For each ticket we have only two possibilities → winning ticket (success)
→ non winning ticket (failure)

probability of a ticket being a winning ticket $p = 0.3$ [because 30% of tickets are winning tickets!]

Buy 8 tickets and we want to know how many are winning tickets

Define r.v. X = # of winning tickets among the 8 tickets purchased

Then $\boxed{X \sim b(\underset{\substack{\uparrow \\ n}}{8},\ \underset{p}{0.3})}$ p ↖ probability of success

n # repetition

Recall: $X \sim bi(n,p)$

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

a) $P(\underbrace{X = 2}_{\text{exactly 2 winning tickets}}) = f(2) = \binom{8}{2} \cdot (0.3)^2 \cdot (0.7)^6 = \ldots\ldots$ calculate

b) $P(\underbrace{X \leq 3}_{\text{at most 3 tickets are winning tickets}}) = \sum_{x=0}^{3} \binom{8}{x} (0.3)^x \cdot (0.7)^{8-x}$ = ------ 4 terms, tedious to evaluate

$= 0.8059$ from table entry corresponding to $\left.\begin{array}{l} n = 8 \\ p = 0.3 \\ x = 3 \end{array}\right\}$

best way use table!!!

c) homework