STA2001 Probability and Statistics (I)

Lecture 5

Tianshi Chen

The Chinese University of Hong Kong, Shenzhen

Review

Definition[Random Variable]

Given a random experiment with sample space S, a function $X : S \to \overline{S} \subseteq R$ that assign one real number $X(s) = x$ to each $s \in S$ is called a Random Variable (RV).

- RV defines a new random experiment with a numeric sample space $\overline{S}$ (take/generate a number from $\overline{S}$)
- If X is one to one, then old random experiment with S $\Leftrightarrow$ new random experiment with $\overline{S}$
- If X is not one to one, then old random experiment with $S \not\Leftrightarrow$ new random experiment with $\overline{S}$
- X is said to be a discrete RV if $\overline{S}$ is finite or countably infinite

Review

Definition[pmf]

Suppose that X is a RV with range $\overline{S}$. Then a function $f(x) : \overline{S} \to (0, 1]$ is called pmf, if

1. $f(x) > 0, \quad x \in \overline{S}.$ $f(x) > 0 , x \in \overline{S}$
2. $\sum_{x\in\overline{S}} f(x) = 1.$ $\sum_{x\in\overline{S}} f(x) = 1$
   $p(X \in A) = \sum_{x\in A} f(x).$
3. $P(X \in A) = \sum_{x\in A} f(x), \quad A \subseteq \overline{S}.$

Note: the 3rd point defines the probability function for an event $A \subseteq \overline{S}$.

The definition domain of $f(x)$ can be extended from $\overline{S}$ to R by simply letting $f(x) = 0$ for $x \notin \overline{S}$.

Review

Definition[cdf]

The function $F(x) : R \to [0,\ 1]$ cdf $F(x) = P(X \le x)$

$= \sum_{x' \le x,\, x' \in \overline{S}} f(x')$

$$F(x) = P(X \le x) = \sum_{x' \le x, x' \in \overline{S}} f(x')$$

is called the cumulative distribution function (cdf).

Definition[Mathematical Expectation]

Assume that X is a discrete RV with range $\overline{S}$ and $f(x)$ is its pmf. If $\sum_{x \in \overline{S}} g(x) f(x)$ exists, then it's called the mathematical expectation of $g(X)$ and is denoted by

期望

$$E[g(X)] = \sum_{x \in \overline{S}} g(x) f(x)$$

Section 2.3 Special Mathematical Expectations [Special $g(X)$]

Mean and Variance

$Var[X] = E[X(X-1)] + E[X] - E^2[X]$

- Mean of a RV $[g(X) = X]$:

$$E[X] = \sum_{x\in\overline{S}} xf(x) \xlongequal{\overline{S}=\{x_1,\cdots,x_k\}} \sum_{i=1}^{k} x_i f(x_i)$$

Interpretation of $E[X]$: the average value of X.

- Variance of a RV $[g(X) = (X - E[X])^2]$:

$$\Downarrow \; !$$

$$Var(X) = E[(X-E[X])^2] = \sum_{x\in\overline{S}} (x-E[X])^2 f(x) = E[X^2]-(E[X])^2$$

$\sum_{x\in\overline{S}} (x^2 - 2x\mu + \mu^2) f(x).$

$\downarrow$

$E[X^2] - 2E[X]\cdot E[X] + E^2[X] = E[X^2] - E^2[X]$

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

- Standard deviation of a RV: the positive square root of the variance, i.e., $\sqrt{Var(X)}$.
- Properties of Variance: Let c be a constant

$$Var(c) = 0, \qquad Var(cX) = c^2 Var(X)$$

Example 1, page 66

Let X equal the number of spots after a 6-sided die is rolled. A reasonable probability model is

$$f(x) = P(X = x) = \frac{1}{6}, \quad x = 1, 2, 3, 4, 5, 6$$

- Mean of X $[g(X) = X]$: Avg value

$$E[X] = \frac{1}{6} \cdot (1 + 2 + 3 + 4 + 5 + 6) = \frac{7}{2}$$

- Variance of X $[g(X) = (X - E[X])^2]$:

$$Var[X] = E[(X - E[X])^2] = E[X^2] - (E[X])^2 = \frac{91}{6} - \frac{49}{4}$$

Variance : Avg of $(X-\mu)^2$

Example 2, page 66 [Interpretation of noise variance and standard deviation]

X has pmf $f(x) = \frac{1}{3}, \quad x = -1, 0, 1$

$$E[X] = 0, \quad Var[X] = \frac{2}{3}, \quad \sigma_X = \sqrt{\frac{2}{3}}$$

Y has pmf $f(y) = \frac{1}{3}, \quad y = -2, 0, 2$

$$E[Y] = 0, \quad Var[Y] = \frac{8}{3}, \quad \sigma_Y = 2\sqrt{\frac{2}{3}}$$

Variance or standard deviation is a measure of the dispersion or

spread out of the values of X with respect to its mean.

The rth Moment

- rth moment of X $[g(X) = X^r$ with r a positive integer]: If $E[X^r] = \sum_{x \in \overline{S}} x^r f(x)$ exists, then it's called the rth moment.

  first moment $E[X]$ second $E[X^2]$

  In addition, if $E[(X-b)^r] = \sum_{x \in \overline{S}} (x-b)^r f(x)$ exists, then it's called the rth moment of X about b, $(X-b)^r$

  and if $E[(X)_r] = E[X(X-1)\cdots(X-r+1)]$ exits, it's called the rth factorial moment. factorial moment

Recall that $\mathrm{Var}[X] = E[X^2] - (E[X])^2$, where $E[X]$ and $E[X^2]$ are the first and second moments, respectively.

Moment Generating Function (mgf)

Definition

Let X be a discrete RV with range space $\overline{S}$ and $f(x)$ be its pmf. If there exists a $h > 0$ such that

$$E[e^{tX}] = \sum_{x \in \overline{S}} e^{tx} f(x) \text{ exists, for } -h < t < h$$

$$E[e^{tX}] = \sum_{X \in \overline{S}} e^{tX} f(X)$$

then the function defined by $M(t) = E[e^{tX}]$ is called the moment generating function (mgf) of X.

The mgf can be used to generate the moments of X.

Properties of Mgf

1. $M(0) = 1$ $M(0)=1$

2. 2 RVs have the same mgf, they have the same probability distribution, i.e., the same pmf.

Same MGF

Same Probability Distribution

Example 3

$$\sum e^{tx} f(x).$$

If X has the mgf

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

then the support of the pmf $f(x)$ of X is $\overline{S} = \{1, 2, 3\}$ and the associated pmf

$$f(x) = \frac{4 - x}{6}, \quad x = 1, 2, 3.$$

$$f(x) = \frac{4-x}{6}$$

Properties of Mgf

3.

$$M'(t) = \sum_{x \in \overline{S}} x e^{tx} f(x)$$
$$M''(t) = \sum_{x \in \overline{S}} x^2 e^{tx} f(x)$$
$$M^{(r)}(t) = \sum_{x \in \overline{S}} x^r e^{tx} f(x)$$

Several questions need to be noted here

- Is $M(t)$ differentiable ? 1st order, 2nd order, $\cdots$, rth order
- Interchange of the differentiation and summation

Properties of Mgf

Setting $t = 0$ leads to

$M'(0) = E[X]$

$$M'(0) = E[X]$$

$M''(0) = E[X^2]$

$$M''(0) = E[X^2]$$

$$M^{(r)}(0) = E[X^r]$$

Observation: the moments can be computed by differentiating

$M(t)$ and evaluating the derivatives at $t = 0$.

Example 4, page 71

Suppose X has the geometric distribution, that is its pmf is

$$f(x) = q^{x-1}p, \quad x = 1, 2, 3, \cdots \quad p = 1 - q, \quad 0 < q < 1$$

Then what is $E(X)$ and $Var(X)$?

Example 4, page 71

几何分布 $f(x) = q^{x-1}p.$ $p = 1 - q$

Suppose X has the geometric distribution, that is its pmf is

$$f(x) = q^{x-1}p, \quad x = 1, 2, 3, \cdots \quad p = 1 - q, \quad 0 < q < 1$$

Then what is $E(X)$ and $Var(X)$? Note the mgf of X is

抛硬币

抛到 x 次 出现了

第一次正面

PMF

$f(x) = P(X = x)$

$= q^{x-1}p.$

$$M(t) = E(e^{tX}) = \sum_{x=1}^{\infty} e^{tx}q^{x-1}p = \frac{p}{q}\sum_{x=1}^{\infty}(qe^t)^x$$

geometric series

$$= (\frac{p}{q})[(qe^t) + (qe^t)^2 + (qe^t)^3 + \cdots]$$

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

$\frac{pe^t}{1-qe^t}$

provided $qe^t < 1$, equivalently $t < -\ln q$

收敛 $\Rightarrow$ $qe^t < 1$ $t < -\ln q$

Example 4, page 71

Let $h = -\ln q$ that is positive. To find the mean and variance of X

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

$$M''(t) = \frac{pe^t(1+qe^t)}{(1-qe^t)^3}$$

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

期望 $E[X] = \frac{1}{p}$

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

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

$Var = \frac{q}{p^2}$

2.4 Binomial Distribution

2项分布

Starting from this section, we will study some typical random phenomena/experiments and corresponding distributions, which are described by RV

1. description of the random phenomena/experiments
2. pmf (probability function), cdf
3. mathematical expectations, e.g., mean, variance, mgf

Bernoulli Experiment

Description: The outcomes can be classified in one of two mutually exclusive and exhaustive ways, say either

success or failure

female or male

life or death

Bernoulli Experiment

Description: The outcomes can be classified in one of two mutually exclusive and exhaustive ways, say either

success or failure

female or male

life or death

Bernoulli Distribution

Let X be a RV associated with a Bernoulli experiment with the probability of success p.

- RV: $X : S \to \overline{S}, S = \{\text{success, failure}\}$. Define

$$X(\text{success}) = 1,\ X(\text{failure}) = 0,\ \overline{S} = \{0, 1\}$$

伯努利分布.

- pmf of $X : f(x) : \overline{S} \to [0, 1]$

$$f(x) = p^x(1-p)^{1-x},\ x \in \overline{S}$$

Then we say X has a Bernoulli distribution with probability of success p.

Bernoulli Distribution

$$(a+b)^n = \sum_{x=0}^{n} \binom{n}{x} b^x a^{n-x}$$

Mathematical expectations:

1. $E[X]$
2. $Var[X]$
3. $M(t) = E[e^{tX}]$

Thus, if we use this binomial expansion with $b = p$ and $a = 1 - p$, then the sum of the binomial probabilities is

$$\sum_{x=0}^{n} \binom{n}{x} p^x (1-p)^{n-x} = [(1-p) + p]^n = 1,$$

a result that had to follow from the fact that $f(x)$ is a pmf.

We now use the binomial expansion to find the mgf for a binomial random variable and then the mean and variance.

The mgf is

$$\begin{aligned} M(t) = E(e^{tX}) &= \sum_{x=0}^{n} e^{tx} \binom{n}{x} p^x (1-p)^{n-x} \\ &= \sum_{x=0}^{n} \binom{n}{x} (pe^t)^x (1-p)^{n-x} \\ &= [(1-p) + pe^t]^n, \qquad -\infty < t < \infty, \end{aligned}$$

$$\begin{aligned} E(e^{tX}) &= \sum_{x=0}^{n} e^{tx} \binom{n}{x} p^x (1-p)^{n-x} \\ &= \sum_{x=0}^{n} \binom{n}{x} [pe^t]^x (1-p)^{n-x} \\ &= [(1-p) + pe^t]^n \end{aligned}$$

$$M'(t) = n[(1-p) + pe^t]^{n-1} \cdot (pe^t)$$

$$M''(t) = n(n-1)[(1-p) + pe^t]^{n-2} (pe^t)^2 + n[(1-p) + pe^t]^{n-1} (pe^t)$$

Bernoulli Distribution

Mathematical expectations:

1. $E[X] = \sum_{x \in \overline{S}} xf(x) = 0 \cdot (1-p) + 1 \cdot p = p$
2. $$Var[X] = E[(X - E[X])^2] = \sum_{x \in \overline{S}} (x-p)^2 f(x)$$
   $$= p^2(1-p) + (1-p)^2 p = (1-p)p$$
3. Mgf: $M(t) = E[e^{tX}] = e^t \cdot p + (1-p), \quad t \in (-\infty, \infty)$

Bernoulli Trials

If a Bernoulli experiment is performed n times

独立性

1. independently, i.e., all trials are independent
2. the probability of success, say p, remains the same from trial to trial.

then these n repetitions of the Bernoulli experiment is called n Bernoulli trials.

Example 1

For a lottery, the probability of winning is 0.001. If you buy the lottery for 10 successive days, that corresponds to 10 Bernoulli trials with the probability of success $p = 0.001$.

Random sample of size n from a Bernoulli distribution

In a sequence of n Bernoulli trials, let X_i denote the Bernoulli RV associated with the ith trial.

An observed sequence of n Bernoulli trials will be n-tuple of zeros and ones, which is called a random sample of size n from a Bernoulli distribution.

Example 2, page 74

Instant lottery ticket; 20% are winners. 5 tickets are purchased and $(0, 0, 0, 1, 0)$ is a random sample. Assuming independence between purchasing different tickets, What is probability of this sample?

Example 2, page 74

Recall that if all trials are independent and let A_i be the event associated with the ith trial. Then

$$P(\cap_{i=1}^{n} A_i) = \prod_{i=1}^{n} P(A_i)$$

Therefore, the probability is $0.2(0.8)^4$ according to multiplication principle for independent events.

Binomial Distribution

We are interested in the number of successes in n Bernoulli trials. The order of the occurrences is not relevant.

Let X be the number of successes in n Bernoulli trials with its range $\overline{S} = \{0, 1, 2, \cdots, n\}$. Find the pmf of X.

1. A Bernoulli (success-failure) experiment is performed n times.
2. The n trials are independent $P(\cap_{i=1}^{n} A_i) = \prod_{i=1}^{n} P(A_i)$, where A_i is the event associated with ith trial, multiplication rule for independent events.
3. The probability of success for each trial is p.

Binomial Distribution

4. If $x \in \overline{S}$ successes occur, the number of ways of selecting x successes in n Bernoulli trials is $\binom{n}{x} = \frac{n!}{x!(n-x)!}$. Since Bernoulli trials are independent, the probability of each way is $p^x(1-p)^{n-x}$

$$\Rightarrow f(x) = P(X = x) = \binom{n}{x} p^x (1-p)^{n-x}, \quad x = 0, 1, \cdots, n$$

Binomial Distribution

Definition[Binomial distribution]

A RV X is said to have a binomial distribution, if the range space $\overline{S} = \{0, 1, \cdots, n\}$ and the pmf

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

and denoted by $X \sim b(n, p)$, where the constants n, p are parameters of the distribution.

It is called the binomial distribution because of its connection with binomial expansion

$$(a+b)^n = \sum_{x=0}^{n} \binom{n}{x} a^x b^{n-x} \text{ with } a = p, \quad b = 1-p$$

Example 2 [revisited]

If X is the number of winning tickets among 5 tickets that are purchased. What is the probability of purchasing 2 winning tickets?

Example 2 [revisited]

If X is the number of winning tickets among 5 tickets that are purchased. What is the probability of purchasing 2 winning tickets?

$$X \sim b(5, 0.2), \quad f(2) = P(X = 2) = \binom{5}{2}(0.2)^2(0.8)^3.$$