STA2001 Probability and Statistics (I)

Lecture 6

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. $f(x) > 0, \quad x \in \overline{S}$.
2. $\sum_{x \in \overline{S}} f(x) = 1$.
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 $F(x) : R \to [0,\ 1]$

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

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

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

Example 1, page 59

Question

Let X be a RV with $\overline{S} = \{-1, 0, 1\}$ and its pmf is $f(x) = \frac{1}{3}$ for $x \in \overline{S}$. What's $E[X^2]$?

Example 1, page 59

Question

Let X be a RV with $\overline{S} = \{-1, 0, 1\}$ and its pmf is $f(x) = \frac{1}{3}$ for $x \in \overline{S}$. What's $E[X^2]$?

$$E[X^2] = \sum_{x \in \overline{S}} x^2 f(x) = (-1)^2 \frac{1}{3} + 0^2 \frac{1}{3} + 1^2 \frac{1}{3} = \frac{2}{3}$$

Theorem 2.2-1, page 60 (Properties of mathematical expectation)

Theorem 2.2-1

Assume that X is a discrete RV with range $\overline{S}$ and $f(x)$ is its pmf. When the involved mathematical expectations exist, the following properties hold:

(a) If c is a constant, $E[c] = c$.

(b) If c is a constant and $g(X)$ is a function.

$$E[cg(X)] = cE[g(X)]$$

(c) If c_1 amd c_2 are constants, $g_1(X)$ and $g_2(X)$ are functions;

$$E[c_1 g_1(X) + c_2 g_2(X)] = c_1 E[g_1(X)] + c_2 E[g_2(X)]$$

Mathematical expectation is a linear operator.

Example 2, page 61

Let $g(X) = (X - b)^2$ where b is a constant to be chosen and suppose $E[(X - b)^2]$ exists. Find the value of b for which $E[(X - b)^2]$ is minimized.

Example 2, page 61

Let $g(X) = (X - b)^2$ where b is a constant to be chosen and suppose $E[(X - b)^2]$ exists. Find the value of b for which $E[(X - b)^2]$ is minimized.

$$\begin{aligned} E[(X-b)^2] &= E[X^2 - 2bX + b^2] \\ &= E[X^2] - 2bE[X] + b^2 \triangleq h(b) \\ \frac{dh(b)}{db} &= -2E[X] + 2b = 0 \qquad \Rightarrow \qquad b = E[X] \end{aligned}$$

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

Mean and Variance

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

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

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

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

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

  - 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

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

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 $[g(X) = X]$:

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

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

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

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.

  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,

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

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

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. $M(0) = 1$
2. 2 RVs have the same mgf, they have the same probability distribution, i.e., the same pmf.

Example 3

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

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

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

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

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

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

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

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

$$\begin{aligned} M(t) = E(e^{tX}) = \sum_{x=1}^{\infty} e^{tx} q^{x-1} p = \frac{p}{q} \sum_{x=1}^{\infty} (qe^t)^x \\ = (\frac{p}{q})[(qe^t) + (qe^t)^2 + (qe^t)^3 + \cdots] \\ = \frac{p}{q} \frac{qe^t}{1 - qe^t} = \frac{pe^t}{1 - qe^t} \end{aligned}$$

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

Example 4, page 71

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

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

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

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

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

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