Chapter 4 part 2

Discrete Random Variables

Jingchen (Monika) Hu

Vassar College

MATH 241

Outline

Variance

- Expected value (or mean) $\mu = E[X]$ yields the weighted average of the possible values of X

Definition

Variance measures the variation (or spread) of these values

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

This holds for all random variable X (not necessary discrete random variable).

- One common simplification:

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

Standard deviation

Definition

Standard Deviation is the square root of the variance

$$\sigma = SD(X) = \sqrt{Var(X)}$$

Question

Let the random variable X denote the GP a certain student will earn in this class. Suppose its pmf is

$$p(0) = 0.05, \quad p(1) = 0.05, \quad p(2) = 0.3, \quad p(3) = 0.4$$

Calculate their GP variance $Var[X]$.

Question

Let the random variable X denote the GP a certain student will earn in this class. Suppose its pmf is

$$p(0) = 0.05, \quad p(1) = 0.05, \quad p(2) = 0.3, \quad p(3) = 0.4$$

Calculate their GP variance $Var[X]$.

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

First compute $p(4) = 1 - p(0) - p(1) - p(2) - p(3) = 0.2$.
Then compute $E[X]$ and $E[X^2]$

$$E[X] = 0 \times 0.05 + 1 \times 0.05 + 2 \times 0.3 + 3 \times 0.4 + 4 \times 0.2 = 2.65$$

$$E[X^2] = 0^2 \times 0.05 + 1^2 \times 0.05 + 2^2 \times 0.3 + 3^2 \times 0.4 + 4^2 \times 0.2 = 8.05$$

Finally we have

$$Var[X] = E(X^2) - (E[X])^2 = 8.05 - 2.65^2 = 1.0275$$

Question

Find $Var(X)$ if

$$P(X = a) = p = 1 - P(X = b)$$

Question

Find $Var(X)$ if
$$P(X = a) = p = 1 - P(X = b)$$

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

First compute $E[X]$ and $E[X^2]$

$$E[X] = a \times p + b \times (1 - p)$$

$$E[X^2] = a^2 \times p + b^2 \times (1 - p)$$

Then we have

$$Var[X] = E(X^2) - (E[X])^2 = a^2 \times p + b^2 \times (1 - p) - (a \times p + b \times (1 - p))^2$$

Properties of variance

$Var(X) \geq 0$

- $Var(X) = 0$ if and only if X is a constant.

If a and b are constants, then

$$Var(aX + b) = a^2 Var(X)$$

- $$Var(aX) = a^2 Var(X)$$
- $$Var(X + b) = Var(X)$$
- $$Var(b) = 0$$

Question

If $E[X] = 1$ and $Var(X) = 5$, find (a) $E[(2 + X)^2]$, (b) $Var(4 + 3X)$.

Question

If $E[X] = 1$ and $Var(X) = 5$, find (a) $E[(2 + X)^2]$, (b) $Var(4 + 3X)$.

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

$$E[aX + b] = aE[X] + b; Var(aX + b) = a^2 Var(X)$$

Part (a):

$$\begin{aligned} E[(2 + X)^2] &= Var(2 + X) + (E[2 + X]^2) \\ &= Var(X) + (2 + E[X])^2 = 5 + 9 = 14 \end{aligned}$$

Part (b)

$$Var(4 + 3X) = 3^2 \times Var(X) = 9 \times 5 = 45$$

Outline

Bernoulli distribution

A trial has two outcomes: success (1) or failure (0).
Let random variable X be the number of success in a single trial.

Definition

Random variable X has a Bernoulli distribution, if

$$P(X = 1) = p, \quad P(X = 0) = 1 - p$$

where $0 \leq p \leq 1$ is the probability of a success.

- The pmf of Bernoulli distribution

$$X \sim \text{Ber}(p) \iff p(1) = p,\ p(0) = 1 - p$$

- Found by Swiss mathematician Jacob Bernoulli.

Recap

Variance σ^2

- **For all random variable**

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

- **Linear function**

$$Var(aX+b) = a^2 Var(X)$$

- **Constants**

$$Var(c) = 0$$

Standard deviation

$$SD(X) = \sqrt{Var(X)}$$

Bernoulli distribution

$$p(1) = p, \quad p(0) = 1-p$$

$$\mu = p, \quad \sigma^2 = p(1-p)$$

Binomial distribution

Definition

Define X to be the number of successes in a fixed number n of independent trials with the same probability of success p as having a Binomial distribution.

Then the pmf of X is $P(X = k) = P($getting k successes in n trials$)$

$$X \sim \textit{Bin}(n, p) \iff p(k) = \binom{n}{k} p^k (1-p)^{n-k}, \quad k = 0, 1, \ldots, n$$

Question

Let X be the number of 6 obtained when roll four fair 6-sided dice simultaneously. Its pmf is.

x	0	1	2	3	4
$p(x)$	0.4823	0.3858	0.1157	0.0154	0.0008

Find its cdf $F(x)$.

Question

Let X be the number of 6 obtained when roll four fair 6-sided dice simultaneously. Its pmf is.

x	0	1	2	3	4
$p(x)$	0.4823	0.3858	0.1157	0.0154	0.0008

Find its cdf $F(x)$.

$$F(x) = \begin{cases} 0 & \text{if } x < 0 \\ 0.4823 & \text{if } 0 \leq x < 1 \\ 0.8681 & \text{if } 1 \leq x < 2 \\ 0.9838 & \text{if } 2 \leq x < 3 \\ 0.9992 & \text{if } 3 \leq x < 4 \\ 1 & \text{if } x \geq 4 \end{cases}$$

Question

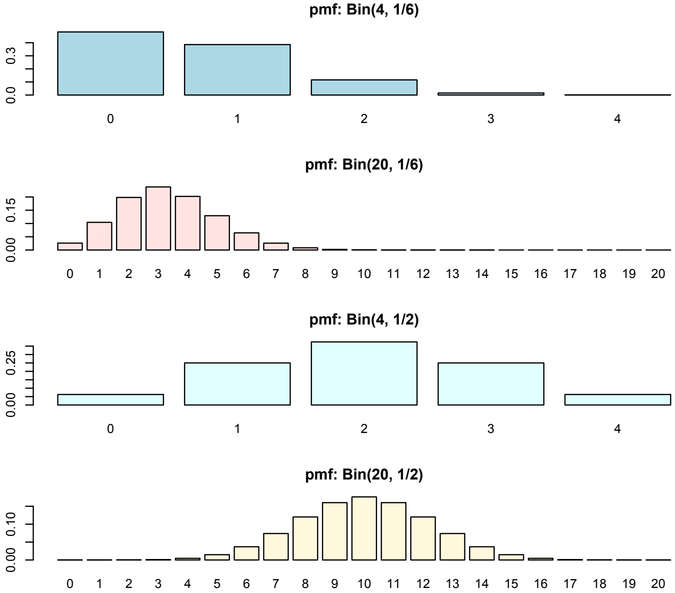
I drink a cup of coffee everyday. About 80% of the time I buy coffee from CraftedKup, about 20% of the time from the Starbucks. Compute the probability I drink 4 cups of Starbucks coffee in 10 days.

Question

I drink a cup of coffee everyday. About 80% of the time I buy coffee from CraftedKup, about 20% of the time from the Starbucks. Compute the probability I drink 4 cups of Starbucks coffee in 10 days.

$$\binom{10}{4} \times 0.2^4 \times 0.8^6 = 210 \times 0.2^4 \times 0.8^6 = 0.088$$

Pmf of Binomial distribution: uni-modal



Binomial pmf is valid (or well-defined)

- Positive: for any $x \in \mathbb{R}$

$$p(x) \geq 0$$

- Total one (required):

$$\sum_{k=0}^{n} p(k) = 1$$

Recall the Binomial Theorem $(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k}$

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

Mean of Binomial random variable

Toss a coin n times, each toss has prob p being a head. On average, total number of heads equals np.

$$E[X] = np$$

Check textbook page 131 for the derivation (not required).

Properties of Binomial distribution

- Variance

$$Var[X] = np(1-p)$$

Check textbook page 132 for the derivation (not required).

- If we have independent Bernoulli random variable's $X_1, X_2, \ldots, X_n$ with the same probability of success p, then their sum has a Binomial distribution

$$X = X_1 + X_2 + \cdots + X_n \sim \text{Bin}(n, p)$$

Recap

Binomial distribution $X \sim \text{Bin}(n, p)$

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

- mean $\mu = np$
- variance $\sigma^2 = np(1-p)$

Question

Random variable $X \sim \text{Binom}(n, p)$, and the value of n is fixed. For any fixed n, find p such that the distribution of X has the largest spread.

Question

Random variable $X \sim \text{Binom}(n, p)$, and the value of n is fixed. For any fixed n, find p such that the distribution of X has the largest spread.

Variance (or standard deviation) is a measurement of spread.

$$\sigma^2 = np(1 - p)$$

Question

Random variable $X \sim \text{Binom}(n, p)$, and the value of n is fixed. For any fixed n, find p such that the distribution of X has the largest spread.

Variance (or standard deviation) is a measurement of spread.

$$\sigma^2 = np(1-p)$$

To find the $\hat{p}$ that maximize the spread, i.e., find the root of the derivative

$$\begin{aligned} \frac{d\sigma^2}{dp} &= n(1 - p + (-p)) = 0 \\ &\Longrightarrow \hat{p} = 1/2 \end{aligned}$$