Tutorial 4

CB2200 Business Statistics, YANG Yihang

September 30, 2019

Intended Learning Outcomes

After this tutorial, you may know how to

- compute expectation and standard deviation from a distribution.
- describe binomial distribution and implement this distribution into specific problems.
- compute probability of specific events of random variables following binomial distribution.
- check whether a random variable following binomial distribution from some events associated with the random variable.

Recap

- Discrete distribution.
- Expectation: $\mu = \mathbb{E}(X) = \sum_{i=1}^{N} x_i P(x = x_i)$
- Variance: $\sigma^2 = \sum_{i=1}^{N} [x_i - \mathbb{E}(X)]^2 P(X = x_i)$
- Standard deviation $= \sqrt{\text{variance}}$
- Binomial distribution
  - 'n' repetition of identical trials
  - 2 mutually exclusive outcomes (success and failure) in each trial
  - Constant probability of success, π, in each trial
  - Trials are independent.
- $P(X = x) = \frac{n!}{x!(n-x)!} \pi^x (1-\pi)^{(n-x)}$
- $\mathbb{E}(X) = n\pi$
- $\sigma = \sqrt{n\pi(1-\pi)}$

Question 1

Q1

In a recent survey concerning the age (to the nearest year) and weight (to the nearest 10 lb) of first-year university students, the following probability distribution was obtained:

Age	Weight				
	100	110	120	130	140
19	0.02	0.09	0.09	0.01	0.02
20	0.06	0.15	α	0.05	0.03
21	0.02	0.06	0.11	0.04	0.05

a) Find the value of α
b) Construct the probability distribution of the weights of these students.
c) A student is selected at random. What do you expect his/her weight to be?
d) What is the standard deviation of the distribution in (b)?
e) Are "age" and "weight" independent? Why or why not?

Question 1

- a) summation of all possible events equals to 1;
- b) Distribution construction
  - 1. confirm object: e.g. weight
  - 2. find the total probability related to the object.
- c)Expectation calculation
- d)standard deviation calculation
- e) Check whether A and B are independent by formula $P(B|A) = P(B)$ or $P(B|A) \neq P(B)$

Q1

a)
$$\alpha = 1 - 0.02 - 0.09 - 0.09 - 0.01 - 0.02 - 0.06 - 0.15 - 0.05 - 0.03 - 0.02 - 0.06 - 0.11 - 0.04 - 0.05 = 0.2$$

b) Probability distribution of weight

Weight	100	110	120	130	140
Probability	0.1	0.3	0.4	0.1	0.1

c) Expected Weight $= 100 \times 0.1 + 110 \times 0.3 + 120 \times 0.4 + 130 \times 0.1 + 140 \times 0.1 = 118$ lb

d) Standard deviation $= \sqrt{(100-118)^2 \times 0.1 + \ldots + (140-118)^2 \times 0.1} = 10.77$lb

e) Take Age= 21 and Weight= 110 lb as example,
Pr (Age=21 and Weight = 110) =0.06
Pr (Age=21) = 0.02+0.06+0.11+0.04+0.05 = 0.28
Pr (Weight=110) = 0.3
Pr (Weight=110) $\times$ Pr (Age=21) $= 0.3 \times 0.28 = 0.084 \neq 0.06$
$\therefore$ Age and Weight are not independent.

Q2

An airline wants to overbook flights in order to reduce the numbers of vacant seats. For a certain flight, it is known that the probabilities of 0, 1, 2 and 3 vacant seats are 0.70, 0.15, 0.10 and 0.05 respectively.

a) Find the mean and standard deviation for the number of vacant seats.
b) What is the expected total number of vacant seats on 100 such flights?

Question 2

Q2

a) Mean $= 0 \times 0.07 + 1 \times 0.15 + 2 \times 0.1 + 3 \times 0.05 = 0.5$

Standard Deviation $= \sqrt{0.7(0-0.5)^2 + 0.15(1-0.5)^2 + 0.1(2-0.5)^2 + 0.05(3-0.5)^2}$

$= 0.8660$

b) Expected total number $= n \times p = 100 \times 0.5 = 50$

Question 3

Q3

Given the following probability distributions:

Distribution A		Distribution B	
X	$P(X)$	X	$P(X)$
0	0.50	0	0.05
1	0.20	1	0.10
2	0.15	2	0.15
3	0.10	3	0.20
4	0.05	4	0.50

a) Compute the expected value for each distribution.
b) Compute the standard deviation for each distribution.
c) Compare the results of distributions A and B.

Question 3

- c) Imagine its histogram for each distribution.

Question 3

Q3

a) Distribution A:

$$\mu = E(X) = \sum_{i=1}^{N} X_i P(X_i)$$
$$= (0)(0.5) + (1)(0.2) + (2)(0.15) + (3)(0.1) + (4)(0.05) = 1$$

Distribution B:

$$\mu = E(X) = \sum_{i=1}^{N} X_i P(X_i) = (0)(0.05) + (1)(0.1) + (2)(0.15) + (3)(0.2) + (4)(0.5) = 3$$

b) Distribution A:

$$\sigma^2 = \sum_{i=1}^{N} \left[X_i - E(X)\right]^2 P(X_i)$$
$$= (0-1)^2(0.5) + (1-1)^2(0.2) + (2-1)^2(0.15) + (3-1)^2(0.1) + (4-1)^2(0.05) = 1.5$$
$$\sigma = 1.2247$$

Distribution B:

$$\sigma^2 = \sum_{i=1}^{N} \left[X_i - E(X)\right]^2 P(X_i)$$
$$= (0-3)^2(0.05) + (1-3)^2(0.1) + (2-3)^2(0.15) + (3-3)^2(0.2) + (4-3)^2(0.5) = 1.5$$
$$\sigma = 1.2247$$

b) Distribution A and B has the same spread but locate at different position. Distribution A is on the left-hand-side of Distribution B.

Q4

You are trying to develop a strategy for investing in two different stocks. The anticipated annual return for a $1,000 investment in each stock has the following probability distribution:

Returns		
Stock X	Stock Y	Probability
-$50	-$100	0.1
20	50	0.3
100	130	0.4
150	200	0.2

a) For each stock, compute the expected return and the standard deviation of return.
b) Do you think that you will invest in stock X or stock Y? Explain.

- Risk averse: lower variance;
- Risk neutral: higher expectation.

Q4

a) Stock X:

expected return = (-50)(0.1)+(20)(0.3)+(100)(0.4)+(150)(0.2)= 71

s.d.of return = $\sqrt{(-50-71)^2(0.1)+\ldots+(150-71)^2(0.2)} = 61.88$

Stock Y:

expected return = (-100)(0.1)+…+(200)(0.2) = 97

s.d. of return = $\sqrt{((-100-97)^2(0.1)+\ldots+(200-97)^2(0.2)} = 84.27$

b) Stock Y gives investor higher expected return than stock X., but also a higher standard deviation. Thus, a risk-averse investor should invest in stock X, while investor who is willing to take a higher risk can expect a higher return from stock Y.

Question 5

Q5

When a customer places an order with Rudy's On-Line Office Supplies, a computerized accounting information system (AIS) automatically checks to see if the customer has exceeded his or her credit limit. Past records indicate that the probability of customers exceeding their credit limit is 0.05. Suppose that, on a given day, 20 customers place orders. Assume that the number of customers that the AIS detects as having exceeded their credit limit is distributed as a binomial random variable.

a) What are the mean and standard deviation of the number of customers exceeding their credit limits?
b) What is the probability that 0 customers will exceed their limits?
c) What is the probability that 1 customer will exceed his or her limit?
d) What is the probability that 2 or more customers will exceed their limits?

Q5

a) X = no. of customers that the AIS detects as having exceeded their credit limit

π = success probability = 0.05

n = 20

$\because$ X is binomial distribution X ~B(n=20, π =0.05)

mean = n π = 20(0.05) = 1

variance = n π (1 - π) = 20(0.05)(0.95) = 0.95

standard deviation = $\sqrt{0.95}$ = 0.9747

b) $$P(X=0) = \frac{20!}{0!(20-0)!}(0.05)^0(0.95)^{20} = 0.3585$$

c) $$P(X=1) = \frac{20!}{1!(20-1)!}(0.05)^1(0.95)^{19} = 0.3774$$

c) $P(X \geq 2) = 1 - P(X=0) - P(X=1) = 1 - 0.3585 - 0.3774 = 0.2642$

Question 6

Q6

For e-commerce merchants, getting a customer to visit a Web site isn't enough. Merchants must also persuade online shoppers to spend money by completing a purchase. Experts at Consumer Consulting estimate that 88% of Web shoppers abandon their virtual shopping carts before completing their transaction. Consider a sample of 20 customers who visit an e-commerce Web site, and assume that the probability that a customer will leave the site before completing the transaction is 0.88. What is the probability that all 20 of the customers will leave the site without completing a transaction?

Question 6

Q6

Let X be the number of customers who will leave the site without completing a transaction.

X~B(20, 0.88)

$P(X=20) = {}_{20}C_{20}(0.88)^{20}(1-0.88)^{0} = 0.0776$

Q7

A task force of CityU sampled 200 students after the mid-term test to ask them whether they went shopping the weekend before the mid-term test or spent the weekend studying, and whether they did well or poorly on the mid-term test. The following result was obtained.

	Did Well on Mid-Term Test	Did Poorly on Mid-Term Test
Studied for Mid-Term Test	90	10
Went Shopping	30	70

a) What is the probability that a randomly selected student did well on the mid-term test or went shopping the weekend before the mid-term test?

b) A random sample of 10 students is selected. What is the probability that 2 of them did well on mid-term test and studied for mid-term test the weekend before the mid-term test? What distribution are you using? Why can you use such distribution?

- Find information from a contingency table.
- Implementation of binomial distribution and the key features of binomial distribution.

Q7

a) P(did well or went shopping) $= \frac{90+70+30}{200} = \frac{190}{200} = 0.95$

b) Binomial distribution is used since

1. no. of trials is fixed
2. two mutually exclusive outcomes
3. independent trials
4. probability of success is constant

Let Y be the no. of students did well on mid-term test and studied for mid-term test the weekend before the mid-term test out of the selected 10 students

$Y \sim B(n = 10, \pi = 0.45)$

$$P(Y=2) = \frac{10!}{2!(10-2)!} 0.45^2 (1-0.45)^{10-2} = 0.0763$$

Question 8

Q8

Suppose that 1,000 patrons of a restaurant were asked whether they preferred beer or wine. 70% said that they preferred beer. 60% of patrons were male. 80% of the males preferred beer.

a) What is the probability a randomly selected patron prefers wine?
b) What is the probability a randomly selected patron is female and prefers wine?
c) Suppose a randomly selected patron prefers wine, what is the probability that the patron is a male?
d) Suppose 5 patrons were selected, what is the probability that at least four of them prefer beer?

Q8

a) Pr(wine) = 1 – 0.7 = 0.3

b) Pr(beer and male) = 0.8 *0.6 = 0.48
=> Pr(wine and male)= 0.6- 0.48 = 0.12
=> Pr(wine and female) = 0.3-0.12 = 0.18

c) Pr(male | wine) = Pr(male and wine) / Pr(wine) = 0.12/0.3 = 0.4

d) Define X be the number of patrons prefer beer in the 5 selected patrons, X~B(5, 0.7) Binomial: n=5, π=0.7

Pr (at least 4 patrons) = Pr (x=4) +Pr(x=5)
= 5!/(4! 1!) $0.7^4 0.3^1$ + 5!/(5! 0!) $0.7^5 0.3^0$
= 0.36015 + 0.16807
= 0.52822

Q9

MTR Corporation has to conduct surveys regularly to evaluate its service quality. According to previous studies, 87% of the passengers refuse to take part in such surveys.

If 15 passengers are selected randomly, what is the probability that at least 2 of them will respond to the survey?

Q9

X is the number of passengers responded to the survey

π is the population proportion of passenger responded to the survey

$\pi = 1 - 0.87 = 0.13$

$X \sim B(15, 0.13)$

$$
\begin{aligned}
\Pr(X \geq 2) &= 1 - \Pr(X=0) - \Pr(X=1) \\
&= 1 - \frac{15!}{0!15!}(0.13)^0(0.87)^{15} - \frac{15!}{1!14!}(0.13)^1(0.87)^{14} \\
&= 1 - 0.1238 - 0.2275 \\
&= 0.5987
\end{aligned}
$$

Q10

According to Dental Association, 60% of all dentists use nitrous oxide ("laughing gas") in their practice. Let x be the number of dentists who use laughing gas in practice in a random sample of five dentists. The probability distribution of x is as follows:

x	0	1	2	3	4	5
P(x)	0.0102	0.0768	0.2304	0.3456	0.2592	0.0778

a) Find the probability that less than 2 dentists use laughing gas in a sample of five.
b) Find E(x). Interpret the result.
c) Find the standard deviation of x.
d) Based on the results of (b) and (c), show that the distribution of x is binomial with $n = 5$ and $\pi = 0.6$.

Question 10

Q10

a) $P(x<2) = 0.0102+0.0768 = 0.087$

b) $E(X) = 0x0.0102+1x0.0768+2x0.2304+3x0.3456+4x0.2592+5x0.0778 = 3.0002$
It means, on average, among 5 dentists 3 of them will use "laughing gas".

c) $V(X) = (0-3)^2x0.0102+(1-3)\ )^2x0.0768+(2-3)\ )^2x0.2304++(3-3)\ )^2x0.3456$
$+(4-3)\ )^2x0.2592+(5-3)\ )^2x0.0778$
$= 1.1998$
Thus, $SD(x) = \sqrt{1.1998} = 1.0954$

d) Define "success" = use laughing gas, "failure" = not use laughing gas, then X represents the number of "success" in n independent trials and the probability of success in each trial is p. According to results of b and c, we have $n\pi = 3$ and $n\pi\ (1-\pi) = 1.1998$. Solve the equations, we have $n = 5$ and $\pi = 0.6$.