The central limit theorem

Worked solutions

Further Maths topic practice

This topic online

1 The number of goals scored in a hockey match in a local league has a Poisson distribution with mean 4.5. A random sample of 60 matches is taken and the mean number of goals per match, $\overline{X}$, is calculated.

a. Write down an approximate distribution for $\overline{X}$, giving its parameters.

b. Explain why the central limit theorem is needed in part (a).

(3 marks)

a) Poisson: mean = variance = λ

$E(X) = 4.5$

$Var(X) = 4.5$

Sample mean: same mean, variance ÷ n

$Var(\bar{X}) = \frac{4.5}{60} = 0.075$

$\bar{X} \approx N(4.5, 0.075)$

b) X is Poisson (discrete), not normal

n = 60 is large

By the CLT, X-bar is approximately normal even though X is not

Free for classroom and home use: print, photocopy and share. Not for resale or re-hosting.

2 A random sample of 120 observations is taken from a Poisson distribution with mean 3.2

Estimate the probability that the mean of the sample is greater than 3.4 (4 marks)

Poisson: mean = variance = 3.2

CLT, n = 120 is large

$$\bar{X} \approx N\left(3.2, \frac{3.2}{120}\right)$$

$$\bar{X} \approx N(3.2, 0.026667)$$

$$sd = \sqrt{0.026667} = 0.163299...$$

Standardise

$$P(\bar{X} > 3.4) = P\left(Z > \frac{(3.4 - 3.2)}{0.163299}\right)$$

$$P(\bar{X} > 3.4) = P(Z > 1.2247)$$

$$P(\bar{X} > 3.4) = 1 - 0.889664...$$

$$\boxed{0.1103}$$

3 A tutor sets a 25-question multiple choice quiz each week. The number of questions a student answers correctly by guessing, X, is modelled by $X \sim B(25, 0.36)$. A random sample of 50 weekly scores $X_1, X_2, ..., X_{50}$ is taken.

Estimate the probability that the mean of the sample is less than 8.6 (4 marks)

Binomial: mean np, variance $np(1-p)$

$$E(X) = 25 \times 0.36 = 9$$

$$Var(X) = 25 \times 0.36 \times 0.64 = 5.76$$

CLT, n = 50

$$\bar{X} \approx N\left(9, \frac{5.76}{50}\right)$$

$$\bar{X} \approx N(9, 0.1152)$$

$$sd = \sqrt{0.1152} = 0.339411...$$

$$P(\bar{X} < 8.6) = P\left(Z < \frac{(8.6 - 9)}{0.339411}\right)$$

$$P(\bar{X} < 8.6) = P(Z < -1.1785)$$

$$\boxed{0.1193}$$

Free for classroom and home use: print, photocopy and share. Not for resale or re-hosting.

4 A random sample of 80 observations is taken from a geometric distribution with parameter 0.25

Estimate the probability that the mean of the sample is less than 3.6 (5 marks)

Geometric: mean $\frac{1}{p}$, variance $\frac{(1-p)}{p^2}$

$$E(X) = \frac{1}{0.25} = 4$$

$$Var(X) = \frac{0.75}{0.25^2} = 12$$

CLT, $n = 80$ is large

$$\bar{X} \approx N\left(4, \frac{12}{80}\right)$$

$$\bar{X} \approx N(4, 0.15)$$

$$sd = \sqrt{0.15} = 0.387298...$$

$$P(\bar{X} < 3.6) = P\left(Z < \frac{(3.6-4)}{0.387298}\right)$$

$$P(\bar{X} < 3.6) = P(Z < -1.0328)$$

$$\boxed{0.1508}$$

Free for classroom and home use: print, photocopy and share. Not for resale or re-hosting.

5 In a fairground game a player throws a ball and scores the points shown on the section it lands in. The number of points scored on one throw, X, has the distribution below.

x	0	1	2	5
$P(X = x)$	0.4	0.25	0.2	0.15

Kwame makes 100 throws and calculates his mean score per throw.

Find an estimate of the probability that his mean score is greater than 1.65 (6 marks)

Mean and variance of one throw

$$E(X) = 0 \times 0.4 + 1 \times 0.25 + 2 \times 0.2 + 5 \times 0.15$$

$$E(X) = 1.4$$

$$E(X^2) = 0 + 0.25 + 0.8 + 3.75 = 4.8$$

$$Var(X) = 4.8 - 1.4^2 = 2.84$$

CLT, $n = 100$

$$\bar{X} \approx N\left(1.4, \frac{2.84}{100}\right)$$

$$\bar{X} \approx N(1.4, 0.0284)$$

$$sd = \sqrt{0.0284} = 0.168523...$$

$$P(\bar{X} > 1.65) = P\left(Z > \frac{(1.65 - 1.4)}{0.168523}\right)$$

$$P(\bar{X} > 1.65) = P(Z > 1.4835)$$

$$P(\bar{X} > 1.65) = 1 - 0.931026...$$

$$\boxed{0.0690}$$

Free for classroom and home use: print, photocopy and share. Not for resale or re-hosting.

6 In a training exercise, each of the 40 archers in a club shoots arrows at a small target until they have hit it 3 times. The probability that an archer hits the target with any one arrow is 0.4, independently of all other arrows. Each archer records the number of arrows they shoot.

Estimate the probability that the mean number of arrows shot by the 40 archers is less than 7 (6 marks)

Shoot until the 3rd hit: negative binomial

$X \sim NB(3, 0.4)$

Mean $\frac{r}{p}$, variance $\frac{r(1-p)}{p^2}$

$$E(X) = \frac{3}{0.4} = 7.5$$

$$Var(X) = \frac{3 \times 0.6}{0.4^2} = 11.25$$

CLT, $n = 40$

$$\bar{X} \approx N\left(7.5, \frac{11.25}{40}\right)$$

$$\bar{X} \approx N(7.5, 0.28125)$$

$$sd = \sqrt{0.28125} = 0.530330...$$

$$P(\bar{X} < 7) = P\left(Z < \frac{(7 - 7.5)}{0.530330}\right)$$

$$P(\bar{X} < 7) = P(Z < -0.9428)$$

$$\boxed{0.1729}$$

Free for classroom and home use: print, photocopy and share. Not for resale or re-hosting.

7 An eight-sided dice has faces numbered 1, 2, 3, 4, 5, 6, 7 and 8. The random variable S is the score when the dice is rolled. Leila rolls the dice 50 times and calculates her mean score.

a. Assuming the dice is fair, find, to 3 significant figures, the value of k such that the probability that her mean score is greater than k is 0.05

b. Explain the relevance of the central limit theorem in part (a).

(7 marks)

a) Discrete uniform on 1, 2, ..., n with n = 8

$$E(S) = \frac{(n+1)}{2} = 4.5$$

$$Var(S) = \frac{(n^2-1)}{12} = \frac{63}{12} = 5.25$$

CLT, 50 rolls

$$\bar{S} \approx N\left(4.5, \frac{5.25}{50}\right)$$

$$\bar{S} \approx N(4.5, 0.105)$$

$$sd = \sqrt{0.105} = 0.324037...$$

Upper 5% point of Z is 1.6449

$$\frac{(k-4.5)}{0.324037} = 1.6449$$

$$k = 4.5 + 1.6449 \times 0.324037$$

$$k = 5.0330...$$

$$\boxed{k = 5.03}$$

b) S is uniform, not normal

n = 50 is large

CLT: the mean score is approximately normal, so the normal distribution can be used

Free for classroom and home use: print, photocopy and share. Not for resale or re-hosting.

8 The number of free tickets, X, won by a player in one round of an arcade game has probability distribution

$$P(X = x) = kx \quad x = 1, 2, 3, 4$$

a. Show that $k = \frac{1}{10}$

b. Find E(X) and Var(X).

c. Ruby plays 64 rounds. Estimate the probability that the mean number of tickets she wins per round is at least 3.2

(7 marks)

a) Probabilities add to 1

$$k + 2k + 3k + 4k = 1$$

$$10k = 1$$

$$\boxed{k = \frac{1}{10}}$$

b)

$$E(X) = \frac{1}{10}(1 + 4 + 9 + 16) = 3$$

$$E(X^2) = \frac{1}{10}(1 + 8 + 27 + 64) = 10$$

$$Var(X) = 10 - 3^2 = 1$$

$$\boxed{E(X) = 3,\ Var(X) = 1}$$

c) CLT, $n = 64$

$$\bar{X} \approx N\left(3, \frac{1}{64}\right)$$

$$sd = \frac{1}{8} = 0.125$$

$$P(\bar{X} \geq 3.2) = P\left(Z \geq \frac{(3.2 - 3)}{0.125}\right)$$

$$P(\bar{X} \geq 3.2) = P(Z \geq 1.6)$$

$$P(\bar{X} \geq 3.2) = 1 - 0.945201...$$

$$\boxed{0.0548}$$

Free for classroom and home use: print, photocopy and share. Not for resale or re-hosting.

9 The number of emails, X, that arrive in an office inbox in a 1-minute period has a Poisson distribution with mean 2. A random sample of n one-minute periods is taken and the sample mean is $\bar{X}$.

Using a suitable approximation, find the smallest value of n for which $P(\bar{X} < 2.2) > 0.95$ (6 marks)

CLT: Poisson mean = variance = 2

$$\bar{X} \approx N\left(2, \frac{2}{n}\right)$$

Standardise and use the 95% point 1.6449

$$P\left(Z < \frac{0.2}{\sqrt{(2/n)}}\right) > 0.95$$

$$\frac{0.2}{\sqrt{(2/n)}} > 1.6449$$

$$0.2\sqrt{n} > 1.6449\sqrt{2}$$

$$\sqrt{n} > 11.631...$$

$$n > 135.28...$$

Check: $n = 135$ gives 0.94983, $n = 136$ gives 0.95045

$$\boxed{n = 136}$$

Free for classroom and home use: print, photocopy and share. Not for resale or re-hosting.

10 The number of loaves, X, bought by a customer at a bakery has the distribution below.

x	1	2	3	4
$\mathbf{P}(X = x)$	0.45	a	b	0.1

$E(X) = 2.05$

a. Find the value of a and the value of b.

b. Show that $\text{Var}(X) = 1.1475$

On a Saturday the bakery expects 200 customers, who buy loaves independently. The bakery bakes 430 loaves.

c. Estimate the probability that the bakery does not have enough loaves for all 200 customers.

d. Find the smallest number of loaves the bakery should bake so that this probability is less than 0.01

(9 marks)

a) Probabilities add to 1

$0.45 + a + b + 0.1 = 1$

$a + b = 0.45$

Use $E(X) = 2.05$

$0.45 + 2a + 3b + 0.4 = 2.05$

$2a + 3b = 1.2$

Subtract $2(a + b) = 0.9$

$b = 0.3$

$a = 0.45 - 0.3 = 0.15$

$\boxed{a = 0.15,\ b = 0.3}$

b)

$E(X^2) = 0.45 + 4(0.15) + 9(0.3) + 16(0.1)$

$E(X^2) = 5.35$

$\text{Var}(X) = 5.35 - 2.05^2$

$\boxed{\text{Var}(X) = 1.1475}$

c) Total loaves T for 200 customers

$E(T) = 200 \times 2.05 = 410$

$\text{Var}(T) = 200 \times 1.1475 = 229.5$

CLT: the total of a large sample is approximately normal

$T \approx N(410, 229.5)$

$sd = \sqrt{229.5} = 15.1493...$

Not enough loaves means $T > 430$

Free for classroom and home use: print, photocopy and share. Not for resale or re-hosting.

$$P(T > 430) = P\left(Z > \frac{(430 - 410)}{15.1493}\right)$$

$$P(T > 430) = P(Z > 1.3202)$$

$$\boxed{0.0934}$$

d) Need $P(T > m) < 0.01$, upper 1% point 2.3263

$$\frac{(m - 410)}{15.1493} = 2.3263$$

$$m = 410 + 2.3263 \times 15.1493$$

$$m = 445.24...$$

$$\boxed{446 \text{ loaves}}$$

Free for classroom and home use: print, photocopy and share. Not for resale or re-hosting.

11 The number of bacteria colonies per 10 ml of water from a reservoir has a Poisson distribution with mean λ. The water company says that λ = 5. An inspector believes that λ is greater than 5. She takes a random sample of 40 measurements and calculates the sample mean, $\overline{X}$.

a. Using the central limit theorem, find the critical region for $\overline{X}$ for a test of H_0: λ = 5 against H_1: λ > 5 at the 5% level of significance. Give the critical value to 3 significant figures.

b. The inspector's sample gives $\bar{x}$ = 5.45. State the conclusion of the test in context.

c. Given that in fact λ = 6, find the power of the test.

d. Describe, with a reason, what would happen to the power of the test if the inspector took a larger sample.

(10 marks)

a) Under H_0, CLT with $n = 40$

$$\text{X-bar} \approx N\left(5, \frac{5}{40}\right)$$

$$\text{X-bar} \approx N(5, 0.125)$$

$$\text{sd} = \sqrt{0.125} = 0.353553...$$

Upper tail, 5%: $z = 1.6449$

$$c = 5 + 1.6449 \times 0.353553$$

$$c = 5.5815...$$

Critical region X-bar > 5.58

b) 5.45 < 5.58: not in the critical region

Do not reject H_0: insufficient evidence that the mean number of colonies is more than 5

c) Power = $P(\text{reject } H_0 \mid \lambda = 6)$

$$\text{X-bar} \approx N\left(6, \frac{6}{40}\right)$$

$$\text{sd} = \sqrt{0.15} = 0.387298...$$

$$P(\text{X-bar} > 5.5815) = P\left(Z > \frac{(5.5815 - 6)}{0.387298}\right)$$

$$P(\text{X-bar} > 5.5815) = P(Z > -1.0804)$$

Power = 0.860

d) $\text{Var(X-bar)} = \frac{\lambda}{n}$ gets smaller as n grows

Power increases: X-bar is more concentrated near 6, so more likely to be in the critical region

Free for classroom and home use: print, photocopy and share. Not for resale or re-hosting.

12 Trainee electricians each repeat a wiring test until they pass it r times. The probability of passing on any attempt is p, independently of other attempts. The number of attempts a trainee needs, Y, has mean 10 and variance 15.

a. Find the value of p and the value of r.

b. A group of 30 trainees take the test. Estimate the probability that the mean number of attempts for the group is greater than 11

c. State one assumption you have made about the trainees in part (b).

(10 marks)

a) Negative binomial: mean $\frac{r}{p}$, variance $\frac{r(1-p)}{p^2}$

$$\frac{r}{p} = 10$$

$$\frac{r(1-p)}{p^2} = 15$$

Divide the second by the first

$$\frac{(1-p)}{p} = 1.5$$

$$1 - p = 1.5p$$

$$p = 0.4$$

$$r = 10 \times 0.4 = 4$$

$p = 0.4, r = 4$

b) CLT, $n = 30$

$$Y\text{-bar} \approx N\left(10, \frac{15}{30}\right)$$

$$Y\text{-bar} \approx N(10, 0.5)$$

$$sd = \sqrt{0.5} = 0.707107...$$

$$P(Y\text{-bar} > 11) = P\left(Z > \frac{(11-10)}{0.707107}\right)$$

$$P(Y\text{-bar} > 11) = P(Z > 1.4142)$$

0.0786

c)

The trainees' numbers of attempts are independent, each with the same p

Free for classroom and home use: print, photocopy and share. Not for resale or re-hosting.