



1

The random variable $X \sim B(200, 0.3)$

- Explain why a normal distribution can be used as an approximation to the distribution of X
- Write down the mean and the variance of this normal distribution.
- Use the normal approximation to estimate $P(X \leq 55)$

(5 marks)

a)

n is large and p is not too close to 0 or 1

b) Mean np , variance $np(1-p)$

$$\mu = 200 \times 0.3 = 60$$

$$\sigma^2 = 200 \times 0.3 \times 0.7 = 42$$

$$Y \sim N(60, 42)$$

c) Continuity correction: $X \leq 55$ includes all of 55

$$P(X \leq 55) \approx P(Y < 55.5)$$

$$P(Y < 55.5) = P\left(Z < \frac{55.5 - 60}{\sqrt{42}}\right)$$

$$P(Y < 55.5) = P(Z < -0.6944)$$

$$P(X \leq 55) \approx 0.2437$$

2

A courier firm claims that its delivery times, in minutes, are normally distributed with mean 45 and standard deviation 8. A manager believes that the mean delivery time has increased. A random sample of 25 deliveries has a mean delivery time of 47.9 minutes.

Assuming the standard deviation is still 8 minutes, test, at the 5% level of significance, whether there is evidence to support the manager's belief. State your hypotheses clearly.

(4 marks)

μ = mean delivery time (minutes)

$$H_0: \mu = 45$$

$$H_1: \mu > 45$$

Under H_0 , the sample mean is normal

$$\bar{X} \sim N\left(45, \frac{8^2}{25}\right)$$

$$\bar{X} \sim N(45, 1.6^2)$$

$$P(\bar{X} \geq 47.9) = 0.0350$$

Compare with the significance level

$$0.0350 < 0.05$$

Significant, so reject H_0

Evidence that the mean delivery time has increased

3

The lengths of a species of fish in a river are normally distributed with mean 28.0 cm and standard deviation 3.2 cm. A biologist believes that the mean length of this species in a nearby lake is different from 28.0 cm. She assumes the standard deviation is also 3.2 cm in the lake. A random sample of 40 fish from the lake has a mean length of 28.9 cm.

- Test, at the 5% level of significance, whether there is evidence to support the biologist's belief. State your hypotheses clearly.
- Find the p -value for this test.

(6 marks)

a) μ = mean length in the lake (cm)

$$H_0: \mu = 28$$

$$H_1: \mu \neq 28$$

Under H_0

$$\bar{X} \sim N\left(28, \frac{3.2^2}{40}\right)$$

Test statistic

$$z = \frac{(28.9 - 28)}{(3.2/\sqrt{40})}$$

$$z = 1.7788$$

Two-tailed 5%: 2.5% in each tail

$$\text{critical values} = \pm 1.96$$

$$1.78 < 1.96$$

Not in the critical region, so do not reject H_0

Insufficient evidence that the mean length in the lake is different

b) Two-tailed, so double the tail

$$P(\bar{X} \geq 28.9) = 0.0376$$

$$p\text{-value} = 2 \times 0.0376$$

$$p\text{-value} = 0.0753$$

4

The lifetime of a type of battery, in hours, is normally distributed with mean μ and standard deviation 1.5. The manufacturer claims that $\mu = 20$. A consumer group suspects that the mean lifetime is less than 20 hours and tests a random sample of 16 batteries.

- Using a 1% level of significance, find the critical region for the sample mean $\bar{X}$.
- The mean lifetime of the 16 batteries is 19.2 hours. Using your answer to part (a), state the conclusion of the test in context.

(5 marks)

$$a) H_0: \mu = 20, H_1: \mu < 20$$

Under H_0

$$\bar{X} \sim N(20, \frac{1.5^2}{16})$$

$$\bar{X} \sim N(20, 0.375^2)$$

Lower tail with 1%

$$P(\bar{X} < c) = 0.01$$

$$\frac{(c - 20)}{0.375} =$$

$$c = 20 - 0.375 \times 2.3263$$

$$c = 19.1276$$

critical region: $\bar{X} < 19.13$

b)

$$19.2 > 19.13$$

Not in the critical region, so do not reject H_0

Insufficient evidence that the mean lifetime is less than 20 hours

5

In a large survey, 45% of adults said that they cycle at least once a week. A random sample of 60 adults is taken and X is the number who cycle at least once a week.

- Find $P(X = 27)$ using a binomial distribution.
- Use a normal approximation to estimate $P(X = 27)$
- Find the percentage error in using the approximation in part (b).
- Use a normal approximation to estimate $P(20 \leq X < 30)$

(6 marks)

$$a) X \sim B(60, 0.45)$$

$$P(X = 27) = {}^{60}C_{27} \times 0.45^{27} \times 0.55^{33}$$

$$P(X = 27) = 0.1031$$

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

$$\mu = 60 \times 0.45 = 27$$

$$\sigma^2 = 60 \times 0.45 \times 0.55 = 14.85$$

$$Y \sim N(27, 14.85)$$

Continuity correction for a single value

$$P(X = 27) \approx P(26.5 < Y < 27.5)$$

$$P(X = 27) \approx 0.1032$$

c)

$$\% \text{ error} = \frac{(0.103236 - 0.103089)}{0.103089} \times 100$$

$$\% \text{ error} = 0.14\%$$

$$d) 20 \leq X < 30 \text{ means } X = 20, 21, \dots, 29$$

$$P(20 \leq X < 30) \approx P(19.5 < Y < 29.5)$$

$$P(20 \leq X < 30) \approx 0.7159$$

6

A garden centre states that 72% of its sweet pea seeds germinate. Seeds germinate independently.

- Darius plants 25 of the seeds. Find the probability that at least 20 of them germinate.
- The garden centre plants 400 of the seeds. Use a suitable approximation to find the probability that more than 300 of them germinate.
- Give a reason why the approximation in part (b) is suitable.

(6 marks)

$$a) X \sim B(25, 0.72)$$

$$P(X \geq 20) = 1 - P(X \leq 19)$$

$$P(X \geq 20) = 1 - 0.7415$$

$$P(X \geq 20) = 0.2585$$

$$b) G \sim B(400, 0.72): \text{approximate with a normal}$$

$$\mu = 400 \times 0.72 = 288$$

$$\sigma^2 = 400 \times 0.72 \times 0.28 = 80.64$$

$$Y \sim N(288, 80.64)$$

More than 300 starts at 301, so use 300.5

$$P(G > 300) \approx P(Y > 300.5)$$

$$P(Y > 300.5) = P(Z > \frac{(300.5 - 288)}{\sqrt{80.64}})$$

$$P(Y > 300.5) = P(Z > 1.3920)$$

$$P(G > 300) \approx 0.0820$$

c)

n is large and p is not close to 0 or 1

7

A machine fills bottles with water. The volume, in ml, in a bottle is normally distributed with mean μ and standard deviation 4. The machine is set so that $\mu = 500$. A quality manager will test $H_0: \mu = 500$ against $H_1: \mu < 500$ at the 5% level of significance, using a random sample of n bottles.

a. The sample mean is 499 ml. Find the least value of n for which this sample mean would lead to H_0 being rejected.

b. State one assumption you have made about the standard deviation.

(5 marks)

a) Under H_0

$$\bar{X} \sim N(500, \frac{4^2}{n})$$

Lower tail 5%: critical $z = -1.6449$

$$z = \frac{(499 - 500)}{(4/\sqrt{n})} < -1.6449$$

$$\frac{-1}{4} < -1.6449$$

$$\frac{\sqrt{n}}{4} > 1.6449$$

$$\sqrt{n} > 6.5794$$

$$n > 43.29$$

n must be a whole number

$$n = 44$$

b)

σ is still 4 ml when the mean changes

8

The resting heart rates, in beats per minute, of adults in a town are normally distributed with mean 72 and standard deviation 9. A sports scientist believes that members of a running club have a lower mean resting heart rate. A random sample of 30 members of the club has a mean resting heart rate of 68.4 beats per minute.

- Assuming that the standard deviation for members of the running club is also 9, test, at the 1% level of significance, whether there is evidence to support the scientist's belief. State your hypotheses and the critical value of your test statistic clearly.
- Explain why the sample mean has a normal distribution.
- State, with a reason, whether the conclusion would be different at the 5% level of significance.

(6 marks)

a) μ = mean resting heart rate of club members

$$H_0: \mu = 72$$

$$H_1: \mu < 72$$

Under H_0

$$\bar{X} \sim N\left(72, \frac{9^2}{30}\right)$$

Test statistic

$$z = \frac{(68.4 - 72)}{(9/\sqrt{30})}$$

$$z =$$

1% lower tail: $P(Z < z) = 0.01$

critical value =

Not in the critical region, so do not reject H_0

Insufficient evidence at 1% of a lower mean heart rate for runners

b)

Heart rates are normal, so the mean of a random sample is normal

c) 5% lower tail critical value = -1.6449

Yes: at 5% H_0 would be rejected

9

The masses of eggs from a farm are normally distributed with mean 65 g and standard deviation 2.5 g. The farmer changes the feed given to the hens and wants to test whether the mean mass of the eggs has changed. She assumes the standard deviation is unchanged and takes a random sample of 20 eggs.

- Write down suitable hypotheses for this test.
- Find the critical region for the sample mean $\bar{X}$, using a 1% level of significance.
- The sample mean is 66.5 g. State the conclusion of the test in context.
- Find the p -value for this test.

(7 marks)

a) μ = mean mass of eggs (g)

$$H_0: \mu = 65$$

$$H_1: \mu \neq 65$$

b) Under H_0

$$\bar{X} \sim N\left(65, \frac{2.5^2}{20}\right)$$

$$\text{standard error} = \frac{2.5}{\sqrt{20}} = 0.5590$$

1% two-tailed: 0.5% in each tail

$$z = \pm 2.5758$$

$$\text{Lower} = 65 - 2.5758 \times 0.5590 = 63.5601$$

$$\text{upper} = 65 + 2.5758 \times 0.5590 = 66.4399$$

$$\bar{X} < 63.56 \text{ or } \bar{X} > 66.44$$

c)

$$66.5 > 66.44$$

In the critical region, so reject H_0

Evidence that the mean mass of the eggs has changed

d) Two-tailed, so double the tail

$$P(\bar{X} \geq 66.5) = 0.0036$$

$$p\text{-value} = 2 \times 0.0036$$

$$p\text{-value} = 0.0073$$

A company makes LED bulbs. Each bulb, independently, has a probability of 0.08 of being faulty. Bulbs are made in large batches. A random sample of 500 bulbs is taken from each batch and the batch is accepted if fewer than 50 of the sample are faulty.

The company aims for at least 95% of batches to be accepted.

- Using a normal approximation, find the probability that a batch is accepted.
- Explain whether the company is likely to achieve its aim.
- Suggest one change to the acceptance rule that would help the company to achieve its aim.

(6 marks)

a) F = number faulty in the sample

$$F \sim B(500, 0.08)$$

$$\mu = 500 \times 0.08 = 40$$

$$\sigma^2 = 500 \times 0.08 \times 0.92 = 36.8$$

$$Y \sim N(40, 36.8)$$

Fewer than 50 means $F \leq 49$, so use 49.5

$$P(F < 50) \approx P(Y < 49.5)$$

$$P(Y < 49.5) = P\left(Z < \frac{(49.5 - 40)}{\sqrt{36.8}}\right)$$

$$P(Y < 49.5) = P(Z < 1.5660)$$

$$P(\text{accept}) \approx 0.9413$$

b) Compare with the target

$$0.9413 < 0.95$$

Unlikely to achieve the aim

c) Check a higher limit

$$P(F < 52) \approx P(Y < 51.5) = 0.9710$$

$$0.9710 > 0.95$$

Accept if fewer than 52 are faulty

Climbing ropes from supplier A have lengths, in metres, that are normally distributed with mean μ and standard deviation 0.5. Given that 10% of these ropes are longer than 25.6 m,

- find the value of μ , giving your answer to 2 decimal places.
- Find the proportion of ropes from supplier A with lengths between 24.5 m and 25.5 m.

Supplier B says that its ropes are longer on average than those from supplier A. The lengths of ropes from supplier B are normally distributed with standard deviation 0.5 m. A random sample of 15 ropes from supplier B has a mean length of 25.21 m.

- Stating your hypotheses clearly and using a 1% level of significance, test whether the mean length of the ropes from supplier B is greater than your value of μ

(9 marks)

a) 10% above 25.6, so 90% below

$$\frac{(25.6 - \mu)}{0.5} = 1.2816$$

$$\mu = 25.6 - 0.5 \times 1.2816$$

$$\mu = 24.9592$$

$$\mu = 24.96$$

b) $L \sim N(24.96, 0.5^2)$

$$P(24.5 < L < 25.5) = 0.6811$$

c) μ = mean length of supplier B's ropes

$$H_0: \mu = 24.96$$

$$H_1: \mu > 24.96$$

Under H_0

$$\bar{X} \sim N(24.96, \frac{0.5^2}{15})$$

$$z = \frac{(25.21 - 24.96)}{(0.5/\sqrt{15})}$$

$$z = 1.9365$$

1% upper tail

$$\text{critical value} = 2.3263$$

$$1.94 < 2.3263$$

Not in the critical region, so do not reject H_0

Insufficient evidence that supplier B's ropes are longer on average

An airline finds that each passenger who books a seat on a flight turns up with probability 0.92, independently of other passengers. A plane has 180 seats and the airline sells 190 tickets for a flight.

- State one reason why the assumption that passengers turn up independently may not be realistic.
- Using a normal approximation, find the probability that more passengers turn up than there are seats.

After a change to its booking rules, the airline believes that the probability of a passenger turning up has decreased. On a random sample of 250 bookings, 222 passengers turned up.

- Using a normal approximation, test at the 5% level of significance whether there is evidence to support the airline's belief. State your hypotheses clearly.

(10 marks)

a)

Groups or families travel together, so they turn up or miss together

b) T = number who turn up

$$T \sim B(190, 0.92)$$

$$\mu = 190 \times 0.92 = 174.8$$

$$\sigma^2 = 190 \times 0.92 \times 0.08 = 13.984$$

$$Y \sim N(174.8, 13.984)$$

More than 180 starts at 181, so use 180.5

$$P(T > 180) \approx P(Y > 180.5)$$

$$P(Y > 180.5) = P\left(Z > \frac{(180.5 - 174.8)}{\sqrt{13.984}}\right)$$

$$P(Y > 180.5) = P(Z > 1.5243)$$

$$P(T > 180) \approx 0.0637$$

c) p = probability a passenger turns up

$$H_0: p = 0.92$$

$$H_1: p < 0.92$$

Under H_0 , $X \sim B(250, 0.92)$

$$\mu = 250 \times 0.92 = 230$$

$$\sigma^2 = 250 \times 0.92 \times 0.08 = 18.4$$

$$Y \sim N(230, 18.4)$$

Continuity correction: $X \leq 222$ becomes $Y < 222.5$

$$P(X \leq 222) \approx P(Y < 222.5)$$

$$P(Y < 222.5) = P(Z < \frac{(222.5 - 230)}{\sqrt{18.4}})$$

$$P(Y < 222.5) = P(Z < -1.7484)$$

$$P(Y < 222.5) = 0.0402$$

$$0.0402 < 0.05$$

Significant, so reject H_0

Evidence that the probability of turning up has decreased