Stats1 Chapter 7: Hypothesis Testing

Chapter Practice

Key Points

1 The area under a continuous probability distribution is equal to 1.

2 If X is a normally distributed random variable, you write $X \sim \mathrm{N}(\mu, \sigma^2)$ where μ is the population mean and σ^2 is the population variance.

3 The normal distribution

- has parameters μ, the population mean, and σ^2, the population variance
- is symmetrical (mean = median = mode)
- has a bell-shaped curve with asymptotes at each end
- has total area under the curve equal to 1
- has points of inflection at $\mu + \sigma$ and $\mu - \sigma$

4 The standard normal distribution has mean 0 and standard deviation 1.

The standard normal variable is written as $Z \sim \mathrm{N}(0, 1^2)$.

5 If n is large and p is close to 0.5, then the binomial distribution $X \sim \mathrm{B}(n, p)$ can be approximated by the normal distribution $\mathrm{N}(\mu, \sigma^2)$ where

- $\mu = np$
- $\sigma = \sqrt{np(1-p)}$

6 If you are using a normal approximation to a binomial distribution, you need to apply a **continuity correction** when calculating probabilities.

7 For a random sample of size n taken from a random variable $X \sim \mathrm{N}(\mu, \sigma^2)$, the sample mean, $\overline{X}$, is normally distributed with $\overline{X} \sim \mathrm{N}\left(\mu, \frac{\sigma^2}{n}\right)$.

8 For the sample mean of a normally distributed random variable, $\overline{X} \sim \mathrm{N}\left(\mu, \frac{\sigma^2}{n}\right)$,

$Z = \dfrac{\overline{X} - \mu}{\frac{\sigma}{\mathrm{R}n}}$ is a normally distributed random variable with $Z \sim \mathrm{N}(0, 1)$.

Chapter Exercises

1 The heights of a large group of men are normally distributed with a mean of 178 cm and a standard deviation of 4 cm. A man is selected at random from this group.

a Find the probability that he is taller than 185 cm. **(2 marks)**

b Find the probability that three men, selected at random, are all less than 180 cm tall. **(3 marks)**

A manufacturer of door frames wants to ensure that fewer than 0.005 men have to stoop to pass through the frame.

c On the basis of this group, find the minimum height of a door frame to the nearest centimetre. **(2 marks)**

2 The weights of steel sheets produced by a factory are known to be normally distributed with mean 32.5 kg and standard deviation 2.2 kg.

a Find the percentage of sheets that weigh less than 30 kg. **(1 mark)**

Bob requires sheets that weigh between 31.6 kg and 34.8 kg.

b Find the percentage of sheets produced that satisfy Bob's requirements. **(3 marks)**

3 The time a mobile phone battery lasts before needing to be recharged is assumed to be normally distributed with a mean of 48 hours and a standard deviation of 8 hours.

a Find the probability that a battery will last for more than 60 hours. **(2 marks)**

b Find the probability that the battery lasts less than 35 hours. **(1 mark)**

A random sample of 30 phone batteries is taken.

c Find the probability that 3 or fewer last less than 35 hours. **(2 marks)**

4 The random variable $X \sim N(24, \sigma^2)$. Given that $P(X > 30) = 0.05$, find:

a the value of σ **(2 marks)**

b $P(X < 20)$ **(1 mark)**

c the value of d so that $P(X > d) = 0.01$. **(2 marks)**

5 A machine dispenses liquid into plastic cups in such a way that the volume of liquid dispensed is normally distributed with a mean of 120 ml. The cups have a capacity of 140 ml and the probability that the machine dispenses too much liquid so that the cup overflows is 0.01.

a Find the standard deviation of the volume of liquid dispensed. **(2 marks)**

b Find the probability that the machine dispenses less than 110 ml. **(1 mark)**

Ten percent of customers complain that the machine has not dispensed enough liquid.

c Find the largest volume of liquid, to the nearest millilitre, that will lead to a complaint. **(2 marks)**

6 The random variable $X \sim N(\mu, \sigma^2)$. The lower quartile of X is 20 and the upper quartile is 40.

a Find μ and σ. **(3 marks)**

b Find the 10% to 90% interpercentile range. **(3 marks)**

7 The heights of seedlings are normally distributed. Given that 10% of the seedlings are taller than 15 cm and 5% are shorter than 4 cm, find the mean and standard deviation of the heights. **(4 marks)**

8 A psychologist gives a student two different tests. The first test has a mean of 80 and a standard deviation of 10 and the student scores 85.

a Find the probability of scoring 85 or more on the first test. **(2 marks)**

The second test has a mean of 100 and a standard deviation of 15. The student scores 105 on the second test.

b Find the probability of a score of 105 or more on the second test. **(2 marks)**

c State, giving a reason, which of the student's two test scores was better. **(2 marks)**

Chapter Exercises

9 Jam is sold in jars and the mean weight of the contents is 108 grams. Only 3% of jars have contents weighing less than 100 grams. Assuming that the weight of jam in a jar is normally distributed, find:

a the standard deviation of the weight of jam in a jar **(2 marks)**

b the proportion of jars where the contents weigh more than 115 grams. **(2 marks)**

A random sample of 25 jars is taken.

c Find the probability that 2 or fewer jars have contents weighing more than 115 grams. **(3 marks)**

10 The waiting time at a doctor's surgery is assumed to be normally distributed with standard deviation of 3.8 minutes. Given that the probability of waiting more than 15 minutes is 0.0446, find:

a the mean waiting time **(2 marks)**

b the probability of waiting less than 5 minutes. **(2 marks)**

11 The thickness of some plastic shelving produced by a factory is normally distributed. As part of the production process the shelving is tested with two gauges. The first gauge is 7 mm thick and 98.61% of the shelving passes through this gauge. The second gauge is 5.2 mm thick and only 1.02% of the shelves pass through this gauge.

Find the mean and standard deviation of the thickness of the shelving. **(4 marks)**

12 A fair coin is spun 60 times. Use a suitable approximation to estimate the probability of obtaining fewer than 25 heads.

Chapter Exercises

13 The owner of a local corner shop calculates that the probability of a customer buying a newspaper is 0.40.

A random sample of 100 customers is recorded.

a Give two reasons why a normal approximation may be used in this situation. **(2 marks)**

b Write down the parameters of the normal distribution used. **(2 marks)**

c Use this approximation to estimate the probability that at least half the customers bought a newspaper. **(2 marks)**

14 The random variable $X \sim B(120, 0.46)$.

a Find $P(X = 65)$. **(1 mark)**

b State why a normal distribution can be used to approximate X, and write down the parameters of such a normal distribution. **(4 marks)**

c Find the percentage error in using the normal approximation to calculate $P(X = 65)$. **(3 marks)**

15 The random variable $Y \sim B(300, 0.6)$.

a Give two reasons why a normal distribution can be used to approximate Y. **(2 marks)**

b Find, using the normal approximation, $P(150 < Y \leqslant 180)$. **(4 marks)**

c Find the largest value of y such that $P(Y < y) < 0.05$. **(3 marks)**

16 Past records from a supermarket show that 40% of people who buy chocolate bars buy the family-size bar. A random sample of 80 people is taken from those who bought chocolate bars. Use a suitable approximation to estimate the probability that more than 30 of these 80 people bought family-size bars.

Chapter Exercises

17 A horticulture company sells apple-tree seedlings. It is claimed that 55% of these seedlings will produce apples within three years.

A random sample of 20 seedlings is taken and X produce apples within three years.

a Find $P(X > 10)$. **(2 marks)**

A second random sample of 200 seedlings is taken. 95 produce apples within three years.

b Assuming the company's claim is correct, use a suitable approximation to find the probability that 95 or fewer seedlings produce apples within three years. **(4 marks)**

c Using your answer to part **b**, comment on the company's claim. **(1 mark)**

18 A herbalist claims that a particular remedy is successful in curing a particular disease in 52% of cases.

A random sample of 25 people who took the remedy is taken.

a Find the probability that more than 12 people in the sample were cured. **(2 marks)**

A second random sample of 300 people was taken and 170 were cured.

b Assuming the herbalist's claim is true, use a suitable approximation to find the probability that at least 170 people were cured. **(4 marks)**

c Using your answer to part **b**, comment on the herbalist's claim. **(1 mark)**

19 The random variable X has a normal distribution with mean μ and standard deviation 2.

A random sample of 25 observations is taken and the sample mean $\overline{X}$ is calculated in order to test the null hypothesis $\mu = 7$ against the alternative hypothesis $\mu > 7$ using a 5% level of significance.

Find the critical region for $\overline{X}$. **(4 marks)**

Chapter Exercises

20 A certain brand of mineral water comes in bottles. The amount of water in a bottle, in millilitres, follows a normal distribution of mean μ and standard deviation 2. The manufacturer claims that μ is 125. In order to maintain standards the manufacturer takes a sample of 15 bottles and calculates the mean amount of water per bottle to be 124.2 millilitres.

Test, at the 5% level, whether or not there is evidence that the value of μ is lower than the manufacturer's claim. State your hypotheses clearly. **(4 marks)**

21 Climbing rope produced by a manufacturer is known to be such that one-metre lengths have breaking strengths that are normally distributed with mean 170.2 kg and standard deviation 10.5 kg. Find, to 3 decimal places, the probability that:

a a one-metre length of rope chosen at random from those produced by the manufacturer will have a breaking strength of 175 kg to the nearest kg **(2 marks)**

b a random sample of 50 one-metre lengths will have a mean breaking strength of more than 172.4 kg. **(3 marks)**

A new component material is added to the ropes being produced. The manufacturer believes that this will increase the mean breaking strength without changing the standard deviation. A random sample of 50 one-metre lengths of the new rope is found to have a mean breaking strength of 172.4 kg.

c Perform a significance test at the 5% level to decide whether this result provides sufficient evidence to confirm the manufacturer's belief that the mean breaking strength is increased. State clearly the null and alternative hypotheses that you are using. **(3 marks)**

22 A machine fills 1 kg packets of sugar. The actual weight of sugar delivered to each packet can be assumed to be normally distributed. The manufacturer requires that,

i the mean weight of the contents of a packet is 1010 g, and

ii 95% of all packets filled by the machine contain between 1000 g and 1020 g of sugar.

a Show that this is equivalent to demanding that the variance of the sampling distribution, to 2 decimal places, is equal to 26.03 g^2. **(3 marks)**

A sample of 8 packets was selected at random from those filled by the machine. The weights, in grams, of the contents of these packets were

1012.6 1017.7 1015.2 1015.7 1020.9 1005.7 1009.9 1011.4

Assuming that the variance of the actual weights is 26.03 g^2,

b test at the 2% significance level (stating clearly the null and alternative hypotheses that you are using) to decide whether this sample provides sufficient evidence to conclude that the machine is not fulfilling condition **i**. **(4 marks)**

23 The diameters of eggs of the little-gull are approximately normally distributed with mean 4.11 cm and standard deviation 0.19 cm.

a Calculate the probability that an egg chosen at random has a diameter between 3.9 cm and 4.5 cm. **(3 marks)**

A sample of 8 little-gull eggs was collected from a particular island and their diameters, in cm, were

4.4, 4.5, 4.1, 3.9, 4.4, 4.6, 4.5, 4.1

b Assuming that the standard deviation of the diameters of eggs from the island is also 0.19 cm, test, at the 1% level, whether the results indicate that the mean diameter of little-gull eggs on this island is different from elsewhere. **(4 marks)**

24 The random variable X is normally distributed with mean μ and variance σ^2.

a Write down the distribution of the sample mean $\overline{X}$ of a random sample of size n. **(1 mark)**

A construction company wishes to determine the mean time taken to drill a fixed number of holes in a metal sheet.

b Determine how large a random sample is needed so that the expert can be 95% certain that the sample mean time will differ from the true mean time by less than 15 seconds. Assume that it is known from previous studies that $\sigma = 40$ seconds. **(4 marks)**

Challenge

A football manager claims to have the support of 48% of all the club's fans.
A random sample of 15 fans is taken.

a Find the probability that more than 8 of these fans supported the manager.

A second random sample of 250 fans was taken, and is used to test the football manager's claim at the 5% significance level.

b Use a suitable approximation to find the critical regions for this test.

It was found that 102 fans said they supported the manager.

c Using your answer to part **b**, comment on the manager's claim.

Chapter Answers

For Chapter 3, student answers may differ slightly from those shown here when calculators are used rather than table values.

1 **a** 0.0401 **b** 0.3307 **c** 188 cm
2 **a** 12.7% or 12.8% **b** 51.1% or 51.2%
3 **a** 0.0668 **b** 0.0521 **c** 0.9314
4 **a** 3.65 **b** 0.1357 **c** 32.5
5 **a** 8.60 ml **b** 0.123 **c** 109 ml
6 **a** $\mu = 30$, $\sigma = 14.8$ or $\sigma = 14.9$ **b** 38.03
7 Mean 10.2 cm, standard deviation 3.76 cm
8 **a** 0.3085
b 0.370 or 0.371
c The first score was better, since fewer of the students got this score or more.
9 **a** 4.25 or 4.26 **b** 0.050 (2 d.p.) **c** 0.8729
10 **a** 8.54 minutes **b** 0.1758
11 Mean 6.12 mm, standard deviation 0.398 mm
12 0.0778
13 **a** n is large and p is close to 0.5.
b $\mu = 40$, $\sigma^2 = 24$
c 0.0262
14 **a** 0.0147
b n is large and p is close to 0.5; $\mu = 55.2$, $\sigma = 5.46$
c 0.68%
15 **a** n is large and p is close to 0.5.
b 0.5232 **c** 166
16 0.6339
17 **a** 0.5914 **b** 0.0197
c Assuming the claim is correct, there would be a less than 2% chance that 95 seedlings produce apples within 3 years. Therefore it is unlikely that the claim is correct.
18 **a** 0.5801 **b** 0.0594
c Assuming the claim is true, there is a less than 6% chance that 170 or more people would be cured out of 300, so it is likely that the herbalist has under-stated the actual cure rate.
19 $\overline{X} > 7.66$ (3 s.f.)
20 Test statistic $= -1.5491\ldots > -1.6449$
Not significant so accept H_0. There is insufficient evidence to suggest that the mean contents of a bottle is lower than the manufacturer's claim.

21 a Accept $0.032 \sim 0.034$ **b** Accept < 0.069

c Test statistic $= 1.4815... < 1.6449$

Not significant so accept H_0. Insufficient evidence of an increase in the mean breaking strength of climbing rope.

22 a $Z = \pm 1.96 \rightarrow 1000 + 1.96\sigma = 1010$

$1.96\sigma = 10 \rightarrow \sigma^2 = 26.03$ (2 d.p)

b Test statistic $= 2.0165... < 2.3263$

Not significant so accept H_0. There is insufficient evidence of a deviation in mean from 1010. So we can assume condition **i** is being met.

23 a Accept $0.845 \sim 0.846$

b Test statistic $= 3.0145... > 2.5758$

Significant so reject H_0. There is evidence that the mean length of eggs from this island is different from elsewhere.

24 a $X \sim N\left(\mu_1 \frac{\sigma^2}{n}\right)$ **b** Need $n = 28$ or more

Challenge

a 0.2510 **b** $\overline{X} \leqslant 105, \overline{X} \geqslant 135$

c 102 is in the critical region, so at the 5% significance level there is evidence to reject the manager's claim. It is probable that less than 48% of people support the manager.