Hypothesis testing for the mean of a large sample

Starter

1. **(Review of last lesson)** A large number of samples of size n is taken from $X \sim \text{Po}(8)$ and approximately $6.2\,\%$ of the sample means are less than 7.5. Estimate the value of n.

Working: $X \sim \text{Po}(8) \quad \Rightarrow \quad E(X) = 8$ and $\text{Var}(X) = 8$ so $\sigma = \sqrt{8}$

Using the CLT, $\bar{X}_n \sim \text{N}\left(8, \frac{8}{n}\right)$

$P(\bar{X}_n < 7.5) = 0.062$: $\quad P(Z < z) = 0.062 \quad \Rightarrow \quad z \approx -1.538$

$z = \dfrac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}}$: $\quad -1.538 = \dfrac{7.5 - 8}{\frac{\sqrt{8}}{\sqrt{n}}}$

$$\sqrt{n} = \frac{\sqrt{8} \times -1.538}{7.5 - 8} \approx 8.700$$

$$n = 75.69$$

The size of the sample is about 76.

2. **(Review of previous material)** Experience has shown that the scores obtained in a particular test are normally distributed with mean score 70 and variance 36. When the test is taken by a random sample of 49 students, the mean score is 68.5. Is there sufficient evidence, at the $4\,\%$ level, that these students have not performed as well as expected?

Working: $X \sim \text{N}(70, 6^2) \quad \Rightarrow \quad \bar{X}_{49} \sim \text{N}\left(70, \frac{6^2}{49}\right)$

$H_0 : \mu = 70$ (the students haven't done worse in the test)
$H_1 : \mu < 70$ (the students have done worse in the test)
$n = 49, \bar{x} = 68.5$, level of significance is $4\,\%$

$p-$value method

$P(\bar{X} < 68.5) = 0.0401 \equiv 4.01\,\%$

Since $4.01\,\% > 4\,\%$, $\bar{x} = 68.5$ ***does not lie*** in the critical region. Therefore, we ***do not reject*** H_0 and conclude that there is evidence to suggest the students have not done worse in the test.

Critical value method

Let x_{cv} be the critical value such that $P(\bar{X} < x_{cv}) = 0.05 \Rightarrow$
$x_{cv} = 68.4994$

Since $\bar{x} = 68.5 \nless 68.4994 = x_{cv}$, we do not reject H_0 and conclude that there is evidence to suggest the students have not done worse in the test.

E.g. 1 A farmer sells eggs claiming their masses have a mean average of 60 g. The supermarket believes the eggs are underweight and collects a random sample of 85 eggs. The sample mean is 58.4 g. Given that the eggs are known to have a variance of 49 g², test the farmer's claim at the $5\,\%$ significance level. State whether you used the central limit theorem and explain your answer.

Working:

$\mu = 60, \sigma^2 = 49$

Since the distribution is not stated but the sample is large enough, the central limit theorem is used to get: $\bar{X}_{85} \sim \text{N}\left(60, \dfrac{49}{85}\right)$

$H_0 : \mu = 60$ g (the mass of the eggs is as the farmer states)
$H_1 : \mu < 60$ (the eggs are underweight)
$n = 85, \bar{x} = 58.4$, level of significance is $5\,\%$

$p-$value method

$P(\bar{X} < 58.4) = 0.0434 \equiv 4.34\,\%$

Since $4.34\,\% < 5\,\%$, $\bar{x} = 58.4$ ***lies*** in the critical region.

Therefore, the supermarket ***rejects*** H_0 and concludes that there is evidence to suggest the eggs are underweight.

Critical value method

Let x_{cv} be the critical value such that $P(\bar{X} < x_{cv}) = 0.05 \Rightarrow$ $x_{cv} = 58.75$

Since $\bar{x} = 58.4 < 58.75 = x_{cv}$, we reject H_0 and conclude that there is evidence to suggest the eggs are underweight.

E.g. 2 A cyclist regularly cycles a route with a mean speed of 25.3 km/h and a standard deviation of 3 km/h. After a period of specific strength training, she rides the same route 30 times and averages 26.1 km/h.

(a) Using the information given, test at the $10\,\%$ significance level whether the strength training has improved the speed of the cyclist.

(b) State whether you used the central limit theorem and explain your answer.

(c) Make one criticism of the hypothesis test and suggest an improvement based on your criticism.

Working:

(a) $\mu = 25.3, \sigma = 3$

Since the distribution is not stated but the sample is large enough, the central limit theorem is used to get: $\bar{X}_{28} \sim \text{N}\left(25.3, \dfrac{3^2}{30}\right)$

$H_0 : \mu = 25.3$ (strength training has not improved average speed)
$H_1 : \mu > 25.3$ (strength training has improved average speed)
$n = 30, \bar{x} = 26.1$, level of significance is $10\,\%$

$p-$value method

$P(\bar{X} < 26.1) = 0.0721 \equiv 7.21\,\%$

Since $7.21\,\% < 10\,\%$, $\bar{x} = 26.1$ ***lies*** in the critical region.

Therefore, we ***reject*** H_0 and conclude that there is evidence to suggest the cyclist has got quicker following strength training.

Critical value method

Let x_{cv} be the critical value such that $P(\bar{X} < x_{cv}) = 0.05 \Rightarrow$ $x_{cv} = 26.0$

Since $\bar{x} = 26.1 > 26.0 = x_{cv}$, we reject H_0 and conclude that there is evidence to suggest the cyclist has got quicker following strength training.

(b) The central limit theorem was used because the distribution of the speeds of the cyclist is not known and because the sample is larger than 25.

(c) The 30 times used in the hypothesis test are not a random sample — they are all the times. If the times were known, a random sample of 26 times could be taken.

E.g. 3 At a supermarket, mince meat is sold in 1.2 kg packs. Trading standards inspectors take a random sample of 14 packs and the masses were measured. They obtained the results that $\sum x_i = 15.764$ and $\sum x_i^2 = 17.9715$.

(a) State an assumption that must be made before a hypothesis test can be carried out.
(b) Test at the $2\,\%$ level whether the packs are significantly underweight.
(c) If the case went to court, what mathematical defence could the supermarket's lawyer use?

Working: (a) The assumption is that the mass of the packs of mince meat follow a normal distribution since the sample is too small to use the central limit theorem.

(b) $\bar{x} = \dfrac{15.764}{14} = 1.126$ kg

Since σ^2 is unknown, it must be calculated from the sample.

$$s^2 = \frac{n}{n-1}\left(\frac{\sum x_i^2}{n} - \bar{x}^2\right)$$
$$= \frac{14}{13}\left(\frac{17.9715}{14} - \left(\frac{15.764}{14}\right)^2\right)$$
$$= 0.01702$$

$X \sim \mathrm{N}(1.2, \sigma^2) \quad \Rightarrow \quad \bar{X}_{14} \sim \mathrm{N}\left(1.2, \dfrac{0.01702}{14}\right)$

$H_0 : \mu = 1.2$ (the mass of mince meat packs are correctly labelled)
$H_1 : \mu < 1.2$ (the mince meat packs are underweight)
$n = 14, \bar{x} = 1.126$, level of significance is $2\,\%$

$p-$value method

$P(\bar{X} < 1.126) = 0.0169 \equiv 1.69\,\%$

Since $1.69\,\% < 2\,\%$, $\bar{x} = 1.126$ ***does lie*** in the critical region. Therefore, we ***reject*** H_0 and conclude that there is evidence to suggest the packs of mince meat are underweight.

Critical value method

Let x_{cv} be the critical value such that $P(\bar{X} < x_{cv}) = 0.02 \Rightarrow$
$x_{cv} = 1.1284$

Since $\bar{x} = 1.1256 < 1.1284 = x_{cv}$, we reject H_0 and conclude that there is evidence to suggest the packs of mince meat are underweight.

(c) The sample is small so that the sample variance will not be close to the variance of the population.

E.g. 4 A machine is set to produce metal rods with length 4.1 cm. After the machine is serviced, the engineer wants to check whether the machine is working correctly and takes a random sample of 30 rods and finds that $\sum x_i = 128.441$ and $\sum x_i^2 = 554.816$. Test at the $5\,\%$ level whether the machine is working correctly using:

(a) the previous variance of the machine, which was 2.7 mm

(b) a new value for the variance based on the sample.

Working: (a) 2.7 mm $\equiv 0.27$ cm

$$\bar{x} = \frac{128.441}{30} = 4.281$$

By the central limit theorem: $\bar{X}_{30} \sim \mathrm{N}\left(4.1, \frac{0.27}{30}\right)$

$H_0 : \mu = 4.1$ (machine is producing rods of the correct length)
$H_1 : \mu \neq 4.1$ (machine is not producing rods of the correct length)
$n = 30, \bar{x} = 4.281$, level of significance is $5\,\%$ but since it is a two-tailed test, the significance level needs to be halved.

$p-$value method

$P(\bar{X} < 4.281) = 0.0282 \equiv 2.82\,\%$

Since $2.82\,\% > 2.5\,\%$, $\bar{x} = 4.281$ ***does not lie*** in the critical region.

Therefore, we ***do not reject*** H_0 and conclude that there is evidence to suggest the machine is producing rods of the correct length.

Critical value method

It's a two-tail test so halve the significance level.

Let x_{cv} be the critical value such that $P(\bar{X} > x_{cv}) = 0.975 \Rightarrow$ $x_{cv} = 4.286$

Since $\bar{x} = 4.281 < 4.286 = x_{cv}$, we do not reject H_0 and conclude that there is evidence to suggest the machine is producing rods of the correct length.

(b)

$$s^2 = \frac{n}{n-1}\left(\frac{\sum x_i^2}{n} - \bar{x}^2\right)$$
$$= \frac{30}{29}\left(\frac{554.816}{30} - \left(\frac{128.441}{30}\right)^2\right)$$
$$= 0.1694$$

By the central limit theorem: $\bar{X}_{30} \sim \mathrm{N}\left(4.1, \frac{0.1694}{30}\right)$

$H_0 : \mu = 4.1$ (machine is producing rods of the correct length)
$H_1 : \mu \neq 4.1$ (machine is not producing rods of the correct length)
$n = 30, \bar{x} = 4.273$, level of significance is $5\,\%$ but since it is a two-tailed test, the significance level needs to be halved.

$p-$value method

$P(\bar{X} < 4.281) = 0.00800 \equiv 0.800\,\%$

Since $0.800\,\% < 2.5\,\%$, $\bar{x} = 4.281$ ***does lie*** in the critical region.

Therefore, we ***reject*** H_0 and conclude that there is evidence to suggest the machine is not producing rods of the correct length.

Critical value method

It's a two-tail test so halve the significance level.

Let x_{cv} be the critical value such that $P(\bar{X} > x_{cv}) = 0.975 \Rightarrow$ $x_{cv} = 4.247$

Since $\bar{x} = 4.281 > 4.247 = x_{cv}$, we reject H_0 and conclude that there is evidence to suggest the machine is not producing rods of the correct length.

Video: Introduction to hypothesis testing using the Central Limit Theorem

Solutions to Starter and E.g.s

Exercise

p169 9A Qu 1i, 2i, 3i, 4-8, (9 red)