

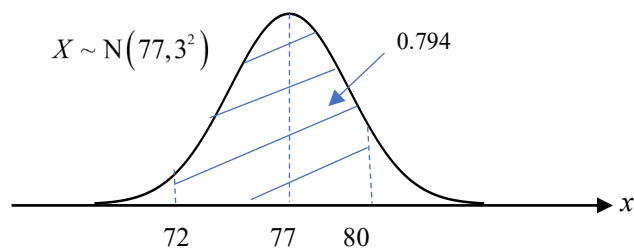
1

- (a) Let X be the random variable denoting the leg length of a man.

$$X \sim N(77, 3^2)$$

$$P(72 < X < 80) = 0.79355 \approx 0.794$$

(b)



(c)

$$\begin{aligned} P(k < X < 72) &= 0.04 \\ P(X < 72) - P(X < k) &= 0.04 \\ P(X < k) &= 0.0077903 \\ k &= 69.744 \\ &= 69.7 \text{ (to 3 s.f.)} \end{aligned}$$

- (d) Let Y be random variable denoting the number of men out of n who will require *Short* trousers.

$$Y \sim B(n, 0.04)$$

$$P(Y \geq 1) \geq 0.9$$

$$1 - P(Y = 0) \geq 0.9$$

$$\binom{n}{0} (0.04)^0 (1 - 0.04)^n \leq 0.1$$

$$(0.96)^n \leq 0.1$$

$$n \lg 0.96 \leq \lg 0.1$$

$$n \geq \frac{\lg 0.1}{\lg 0.96}$$

$$n \geq 56.41$$

$$\therefore \min n = 57$$

$$H_0 : \mu = 77$$

$H_1 : \mu > 77$, where μ is the population mean leg length of a man

Under H_0 ,

$$X \sim N(77, 3^2)$$

$$\bar{X} \sim N\left(77, \frac{3^2}{50}\right)$$

$$Z = \frac{\bar{X} - 77}{\sqrt{\sigma^2/n}} \sim N(0, 1)$$

At 1% level of significance, we reject H_0 if $p\text{-value} \leq 0.01$.

$$p\text{-value} = 0.00109$$

Since $p\text{-value} \leq 0.01$, we reject H_0 and conclude that there is sufficient evidence at the 1% level of significance to claim that the mean men's leg length is more than 77 cm.

The manager needs to assume that the population variance of the men's mean leg length remains unchanged.



2

(a) Unbiased estimates of the population mean

$$\begin{aligned} &= \frac{\sum (x - 650)}{n} + 650 \\ &= -\frac{34.39}{50} + 650 \\ &= 649.3122 \end{aligned}$$

Unbiased estimates of the population variance, s^2

$$\begin{aligned} &= \frac{1}{n-1} \left(\sum (x - 650)^2 - \frac{(\sum (x - 650))^2}{n} \right) \\ &= \frac{1}{49} \left(22769.98 - \frac{(-34.39)^2}{50} \right) \\ &= \frac{89621}{6087} \text{ or } 464.21 \\ &= 464 \text{ (to 3 s.f.)} \end{aligned}$$

(b) $H_0 : \mu = 650$

$H_1 : \mu < 650$, μ is the population mean travelling distance on a single charge

Under H_0 ,

Since $n = 50$ is large,

$$\begin{aligned} \bar{X} &\sim N\left(650, \frac{464.21}{50}\right) \text{ approximately by CLT} \\ Z &= \frac{\bar{X} - 650}{\sqrt{s^2/n}} \sim N(0, 1) \end{aligned}$$

At 5% level of significance, we reject H_0 if $p\text{-value} \leq 0.05$.

$$p\text{-value} = 0.411$$

Since $p\text{-value} > 0.05$, we do not reject H_0 and conclude that there is insufficient evidence at the 5% level of significance that the car manufacturer has overstated the travelling distance on a single charge.

- (c) The MeTube car reviewer **needs** to apply Central Limit Theorem (CLT) because the **distribution of the travelling distance on a single charge is unknown (i.e. not normally distributed)**.

With the sample size of 50 (large), CLT can be applied and the mean travelling distance on a single charge is approximately normally distributed.

- (d) The TokTik car reviewer should use a 2-tail test since travelling distance on a single charge can either be **more than or less than 650 km**.
- (e) The TokTik car reviewer needs to assume that the travelling distance on a single charge is normally distributed.

He also needs to assume that the observations of travelling distance on a single charge are independent.



3

(a) Let X be the random variable denoting the wait time for a customer, in minutes.

Using G.C.,

unbiased estimate of population mean = $3.0556 = 3.06$ (to 3 s.f.)

unbiased estimate of population variance, $s^2 = 0.46667^2 = 0.21778 = 0.218$ (to 3 s.f.)

(b)(i) $H_0 : \mu = 3.3$

$H_1 : \mu < 3.3$, where μ is the population mean wait time for a customer.

Under H_0 ,

$$\begin{aligned}\bar{X} &\sim N\left(3.3, \frac{0.22}{9}\right) \\ Z &= \frac{\bar{X} - 3.3}{\sqrt{\frac{\sigma^2}{9}}} \sim N(0,1)\end{aligned}$$

At 5% level of significance, we reject H_0 if $p\text{-value} \leq 0.05$.

$p\text{-value} = 0.0590$

Since $p\text{-value} > 0.05$, we do not reject H_0 and conclude that there is insufficient evidence at 5% significance level to claim that the mean waiting time has improved.

Assumption:

The customers' waiting time follows a normal distribution.

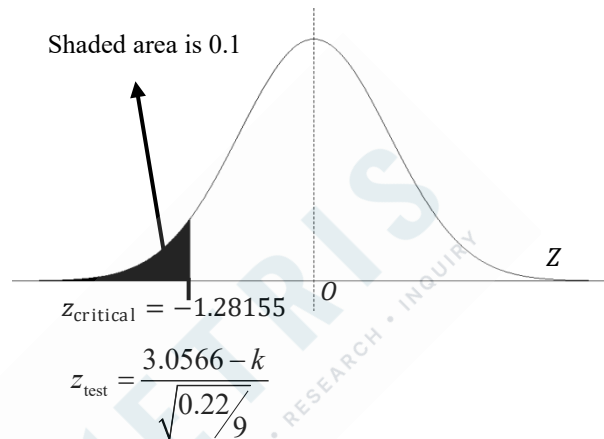
(b)(ii) The p -value is the probability of getting a sample of 9 customers whose wait time is less than or equal to 3.06 minutes when the population mean is 3.3 minutes is 0.0590.

(c) $H_0 : \mu = k$
 $H_1 : \mu < k$

Under H_0 , $\bar{X} \sim N\left(k, \frac{0.22}{9}\right)$

$$Z = \frac{\bar{X} - k}{\sqrt{\sigma^2/n}} \sim N(0,1)$$

At 10% level of significance,



For the manager's claim to be justified, we reject H_0 ,

$$\frac{3.0566 - k}{\sqrt{0.22/9}} \leq -1.28155$$

$$k \geq 3.2559$$

$$k \geq 3.26 \quad (\text{to 3 s.f.})$$

4

(a) Let X be the random variable denoting the volume of drink dispensed in a cup
Sample mean, $\bar{x}=196.375$

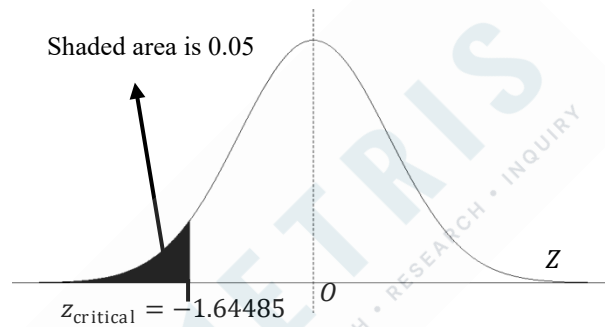
(b) $H_0 : \mu = \beta$
 $H_1 : \mu < \beta$, where μ is the population mean volume of drink dispensed in a cup.

(c) Under H_0 ,

$$\bar{X} \sim N\left(\beta, \frac{6.3^2}{8}\right)$$

$$Z = \frac{\bar{X} - \beta}{\sqrt{\sigma^2/n}} \sim N(0,1)$$

At 5% level of significance,



$$z_{\text{test}} = \frac{196.375 - \beta}{\sqrt{6.3^2/8}}$$

To conclude that the vending machine is dispensing less than β , we reject H_0 ,

$$\frac{196.375 - \beta}{\sqrt{6.3^2/8}} \leq -1.64485$$

$$\beta \geq 200.039$$

$$\beta \geq 201 \text{ (to 3 s.f.)}$$

(d) Since the volume of water after repair **is not known to be normally distributed**, the manager has to take a larger sample (with sample size of at least 30) so that the **sample mean volume of drink** can be approximated to a normal distribution, using Central Limit Theorem.

5

(i) Each scented candle has an **equal chance** of being **selected**. The **selection** of the scented candles is **independent** of each other.

(ii) Unbiased estimates of the population mean

$$\begin{aligned} &= \frac{69}{30} + 600 \\ &= 602.3 \end{aligned}$$

Unbiased estimates of the population variance, s^2

$$\begin{aligned} &= \frac{1}{29} \left(859 - \frac{69^2}{30} \right) \\ &= 24.148 \\ &= 24.1 \text{ (to 3 s.f.)} \end{aligned}$$

(iii) Let X be the random variable denoting the mass, in grams, of a scented candle.

$$H_0 : \mu = 600$$

$$H_1 : \mu \neq 600, \text{ where } \mu \text{ is the population mean mass of a scented candle.}$$

Under H_0 ,

Since $n = 30$ is large,

$$\bar{X} \sim N\left(600, \frac{24.148}{30}\right) \text{ approximately by CLT}$$

$$Z = \frac{X - 600}{\sqrt{s^2/n}} \sim N(0, 1)$$

At 1% level of significance, we reject H_0 if $p\text{-value} \leq 0.01$.

$$p\text{-value} = 0.0104$$

Since $p\text{-value} > 0.01$, we do not reject H_0 and conclude that there is insufficient evidence at the 1% level of significance to claim that the population mean mass of scented candles is not 600 grams.

(iv) The $p\text{-value}$ of 0.0104 is the probability of getting a sample of 30 scented candles whose mass is 602.3 grams when the population mean mass is 600 grams.

(v) $H_0 : \mu = 600$

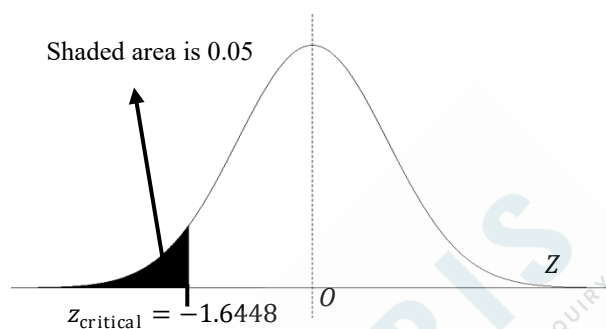
$H_1 : \mu < 600$

Under H_0 ,

$$\bar{X} \sim N\left(600, \frac{25}{n}\right).$$

$$Z = \frac{\bar{X} - 600}{\sqrt{\sigma^2/n}} \sim N(0,1)$$

At 5% level of significance,



$$z_{\text{test}} = \frac{598.2 - 600}{\sqrt{25/n}}$$

For the mass of the scented candles to be overstated, we reject H_0 ,

$$\frac{598.2 - 600}{\sqrt{25/n}} \leq -1.6448$$

$$\sqrt{25/n} \leq 1.094358$$

$$\sqrt{n} \geq 4.5689222$$

$$n \geq 20.875$$

$$n \geq 20.9 \text{ (to 3 s.f.)}$$

6

- (a) Tim can **randomly select 20 Year 1 Boys, 20 Year 1 Girls, 20 Year 2 Boys and 20 Year 2 Girls** to ensure that the sample of students is unbiased and representative of the cohort.
- (b) Tim should conduct a one-tailed test as he believes that students are spending more than 3 hours each day on their homework.

(c) Unbiased estimate of population mean $= \frac{16}{80} + 3$
 $= 3.2$

Unbiased estimate of population variance is $s^2 = \frac{1}{79} \left(164 - \frac{16^2}{80} \right)$
 $= \frac{804}{395}$ or $2.0354430 = 2.04$

- (d) Let X be the random variable denoting the time that a student spends on homework each day (in hours).

$$H_0 : \mu = 3$$

$$H_1 : \mu > 3, \text{ where } \mu \text{ is the population mean time that a students spend on homework each day}$$

Under H_0 ,

Since $n = 80$ is large,

$$\bar{X} \sim N\left(3, \frac{2.0354}{80}\right) \text{ approximately by CLT}$$

$$Z = \frac{\bar{X} - 3}{\sqrt{s^2/n}} \sim N(0, 1)$$

At 5% level of significance, we reject H_0 if $p\text{-value} \leq 0.05$.

$$p\text{-value} = 0.105$$

Since $p\text{-value} > 0.05$, we do not reject H_0 and conclude that there is insufficient evidence at 5% significance level to claim that the population mean time that students spend on homework each day is more than 3 hours. Thus Tim's belief is not accurate at 5% significance level.

$$H_0 : \mu = 3$$

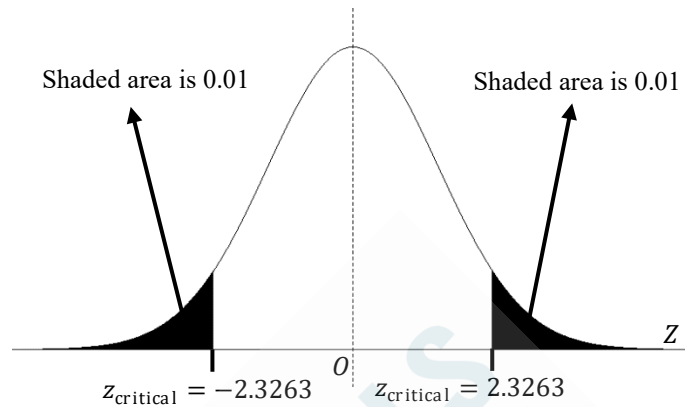
$$H_1 : \mu \neq 3$$

Under H_0 ,

$$\bar{X} \sim N\left(3, \frac{\sigma^2}{10}\right).$$

$$Z = \frac{\bar{X} - 3}{\sqrt{\sigma^2/n}} \sim N(0,1)$$

At 2% level of significance,



$$z_{\text{test}} = \frac{3.2 - 3}{\sqrt{\sigma^2/10}}$$

Given that the teachers' claim was rejected, we reject H_0 ,

$$\frac{3.2 - 3}{\sqrt{\sigma^2/10}} \leq -2.3263 \quad \text{or} \quad \frac{3.2 - 3}{\sqrt{\sigma^2/10}} \geq 2.3263$$

$$\sigma < -0.27187 \quad \text{or} \quad \sigma < 0.27187 \quad (3 \text{ s.f.})$$

(rejected)

$$\therefore 0 < \sigma < 0.27187 \quad (\text{since } \sigma > 0)$$

Assumption:

The time students spending on their homework in a day follows a normal distribution.

7

- (i) Let X be the random variable denoting the mass of a badminton racket.

$$H_0: \mu = 800$$

$$H_1: \mu \neq 800, \text{ where } \mu \text{ is the population mean mass of a badminton racket.}$$

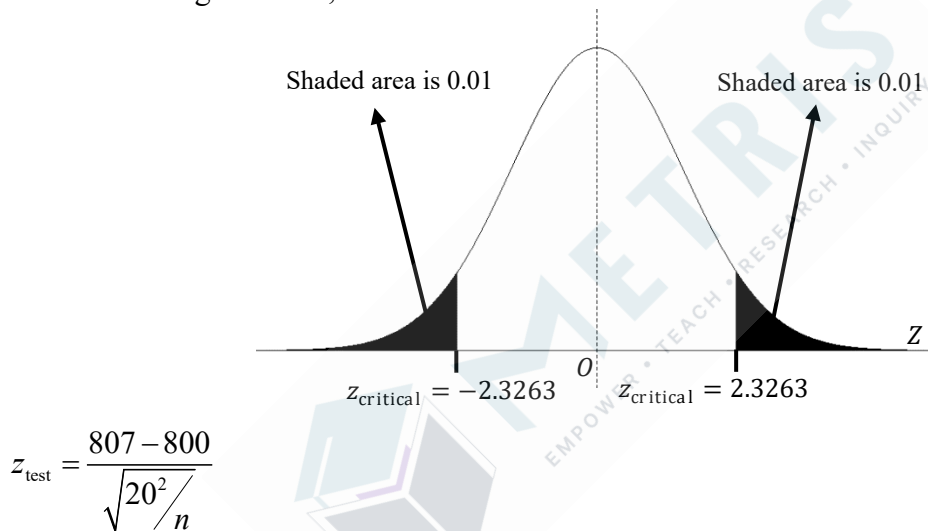
Under H_0 ,

Since n is large,

$$\bar{X} \sim N\left(800, \frac{20^2}{n}\right) \text{ approximately by CLT}$$

$$Z = \frac{\bar{X} - 800}{\sqrt{\sigma^2/n}} \sim N(0, 1)$$

At 2% level of significance,



To conclude that the mean of badminton rackets is not 800g, we reject H_0 ,

$$\begin{aligned} \frac{807 - 800}{\sqrt{20^2/n}} &\leq -2.3263 \quad \text{or} \quad \frac{807 - 800}{\sqrt{20^2/n}} \geq 2.3263 \\ \frac{7\sqrt{n}}{20} &\leq -2.3263 \quad \text{or} \quad \frac{7\sqrt{n}}{20} \geq 2.3263 \\ \text{(no real solution)} & \quad \quad \quad \sqrt{n} \geq 6.6467 \\ & \quad \quad \quad n \geq 44.179 \end{aligned}$$

$\therefore$ least $n = 45$

- (ii) There is no need to know anything about the distribution of the population since the sample size is large, by Central Limit Theorem, the sample mean distribution is approximately normal.

8

(i) Unbiased estimate of population mean $= \frac{\sum x}{n} = \frac{10446}{30} = 348.2$ (exact)

Unbiased estimate of population variance, s^2

$$= \frac{1}{n-1} \left(\sum x^2 - \frac{(\sum x)^2}{n} \right) = \frac{1}{29} \left(3638000 - \frac{(10446)^2}{30} \right) = 24.23448$$

$$H_0 : \mu = 350$$

$H_1 : \mu < 350$, where μ is the population mean number of calories in a Perfect Protein bar.

Under H_0 ,

Since $n = 30$ is large,

$$\bar{X} \sim N \left(350, \frac{24.23448}{30} \right) \text{ approximately by CLT}$$

$$Z = \frac{\bar{X} - 350}{\sqrt{s^2/n}} \sim N(0, 1)$$

At 3% level of significance, we reject H_0 if $p\text{-value} \leq 0.03$.

$$p\text{-value} = 0.022604585 = 0.0226 \text{ (to 3 s.f.)}$$

Since $p\text{-value} \leq 0.03$, we reject H_0 and conclude that there is sufficient evidence at the 3% level of significance to claim that the mean number of calories in Perfect Protein bars is less than 350 cal.

(ii) The test would be inadmissible if the nutritionist had taken a random sample of 15 energy bars as the distribution of the population is unknown (and Central Limit Theorem cannot be applied). The nutritionist will need to assume that the number of calories in Perfect Protein bars follow a normal distribution.

(iii) $H_0 : \mu = 350$

$$H_1 : \mu \neq 350$$

At $\alpha\%$ level of significance, we reject H_0 if $p\text{-value} \leq \frac{\alpha}{100}$.

$$p\text{-value} = 2(0.022604585) = 0.04520917$$

To reject H_0 , $0.04520917 \leq \frac{\alpha}{100}$

$$\alpha \geq 4.520917$$

$$\alpha \geq 4.53 \text{ (to 3 s.f.)}$$

(iv) $H_0 : \mu = 350$

$H_1 : \mu > 350$, where μ is the population mean number of calories in a Dynamic Protein bar.

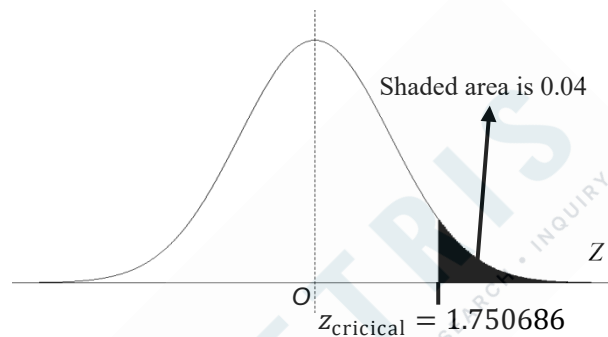
Under H_0 ,

Since n is large,

$$\bar{X} \sim N\left(350, \frac{\sigma^2}{n}\right) \text{ approximately by CLT}$$

$$Z = \frac{\bar{X} - 350}{\sqrt{\sigma^2/n}} \sim N(0, 1)$$

At 4% level of significance, we reject H_0 if $p\text{-value} \leq 0.04$.



$$z_{\text{test}} = \frac{352.2 - 350}{\sqrt{\sigma^2/45}}$$

Since there is no reason to believe that the mean number of calories in the new product is more than 350 calories, we do not reject,

$$\frac{352.2 - 350}{\sqrt{\sigma^2/45}} < 1.750686$$

$$1.750686 \sqrt{\sigma^2/45} > 2.2$$

$$\sigma^2 > 45 \left(\frac{2.2}{1.750686} \right)^2$$

$$\sigma^2 > 71.0626$$

$$\therefore \sigma^2 > 71.1 \text{ (to 3 s.f.)}$$

9

- (a) Every sausage roll produced by the machine has an equal probability of being selected in the sample and each selection is independent of one another.
- (b) Since the lengths of the sausage rolls are around 10cm, (or lengths of the sausage rolls are designed/targeted to be 10cm), by subtracting, the values are smaller and easier to calculate.

- (c) Unbiased estimate of population mean

$$\begin{aligned} &= \frac{-2.5}{60} + 10 \\ &= 9.958333 \\ &= 9.96 \text{ (to 3 s.f.)} \end{aligned}$$

Unbiased estimate of population variance, s^2

$$\begin{aligned} &= \frac{1}{59} \left[4.3 - \frac{(-2.5)^2}{60} \right] \\ &= 0.0711158192 \\ &= 0.0711 \text{ (to 3 s.f.)} \end{aligned}$$



METRIS
EMPOWER • TEACH • RESEARCH • INQUIRY

- (d) $H_0 : \mu = 10$
 $H_1 : \mu < 10$, where μ is the population mean length of a sausage roll

Under H_0 ,

Since $n = 60$ is large,

$$\bar{X} \sim N\left(10, \frac{0.0711158192}{60}\right) \text{ approximately by CLT}$$

$$Z = \frac{\bar{X} - 10}{\sqrt{s^2/n}} \sim N(0, 1)$$

At 10% level of significance, we reject H_0 if $p\text{-value} \leq 0.10$.

$$p\text{-value} = 0.113$$

Since $p\text{-value} > 0.10$, we do not reject H_0 and conclude that there is insufficient evidence at the 10% level of significance to conclude that the manager's claim is valid.

$$H_0 : \mu = 10$$

$$H_1 : \mu \neq 10, \text{ where } \mu \text{ is the population mean length of a sausage roll}$$

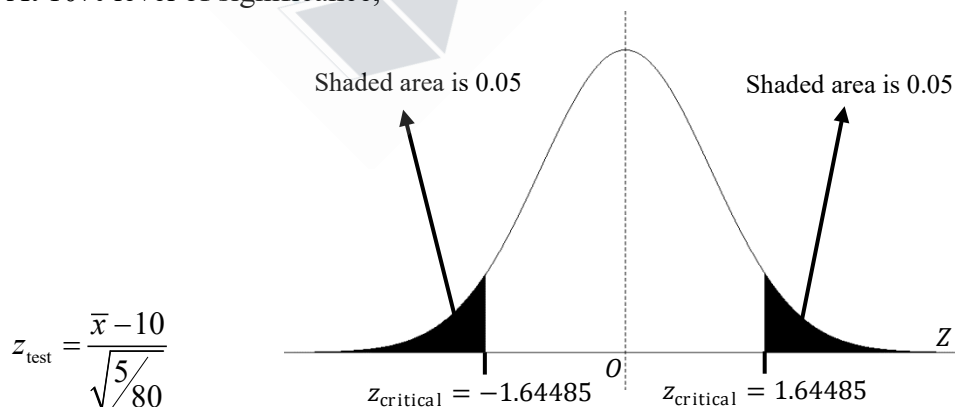
Under H_0 ,

Since $n = 80$ is large,

$$\bar{X} \sim N\left(10, \frac{5}{80}\right) \text{ approximately by CLT}$$

$$Z = \frac{\bar{X} - 10}{\sqrt{\sigma^2/n}} \sim N(0, 1)$$

At 10% level of significance,



Since there is no reason to reject the null hypothesis,

$$-1.64485 \leq \frac{\bar{x} - 10}{\sqrt{5/80}} \leq 1.64485$$

$$-0.4112125 < \bar{x} - 10 < 0.4112125$$

$$9.5887875 < \bar{x} < 10.4112125$$

$$9.59 < \bar{x} < 10.4$$

10

- (a) Since not every student use the social media or follow Mr Kim on the social media platform, hence not every students of Junior colleges have an equal chance of being selected.

- (b) Unbiased estimate of population mean

$$\begin{aligned} &= \frac{\sum x}{n} \\ &= \frac{1707}{50} \\ &= 34.14 \end{aligned}$$

Unbiased estimate of population variance s^2

$$\begin{aligned} &= \frac{1}{n-1} \left[\sum x^2 - \frac{(\sum x)^2}{n} \right] \\ &= \frac{1}{50-1} \left[71695 - \frac{(1707)^2}{50} \right] \\ &= 273.84 \\ &= 274 \text{ minutes}^2 \end{aligned}$$

- (c) $H_0 : \mu = 37$

$H_1 : \mu < 37$, where μ is the population mean time JC student student spent watching Kdrama in a day in minutes.

Under H_0 ,

Since $n = 50$ is large,

$$\bar{X} \sim N\left(37, \frac{273.84}{50}\right) \text{ approximately by CLT}$$

$$Z = \frac{\bar{X} - 37}{\sqrt{s^2/n}} \sim N(0, 1)$$

At 5% level of significance, we reject H_0 if $p\text{-value} \leq 0.05$.

$$p\text{-value} = 0.111$$

Since $p\text{-value} > 0.05$, we do not reject H_0 and conclude that there is insufficient evidence at the 5% level of significance to claim that Ms Shin has overstated the average time spent.

- (d) “5% level of significance” refers to the probability of 0.05 that we claim that the mean time JC student spent watching Kdrama in a day is less than 37 minutes, when is actually 37 minutes.

- (e) $H_0 : \mu = 37$
 $H_1 : \mu \neq 37$, where μ is the population mean time JC student student spent watching Kdrama in a day in minutes.

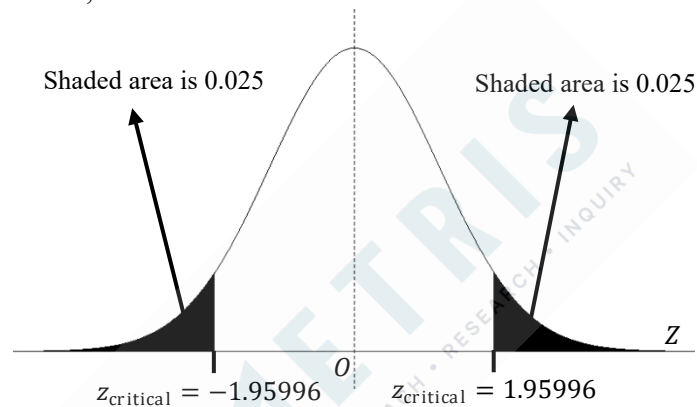
Under H_0 ,

$$X \sim N(37, 285)$$

$$\bar{X} \sim N\left(37, \frac{285}{50}\right)$$

$$Z = \frac{\bar{X} - 37}{\sqrt{\sigma^2/n}} \sim N(0, 1)$$

At 5% level of significance,



$$z_{\text{test}} = \frac{k - 37}{\sqrt{285/50}}$$

Since there is sufficient evidence to conclude that the average time spent watching Kdrama is different from what Ms. Shin claimed, we reject H_0 .

$$\frac{k - 37}{\sqrt{285/50}} \leq -1.95996 \quad \text{or} \quad \frac{k - 37}{\sqrt{285/50}} \geq 1.95996$$

$$k \leq 32.32066 \quad \text{or} \quad k \geq 41.67934$$

$$\therefore 0 \leq k < 32.3 \quad \text{or} \quad 41.7 < k \leq 1440$$

11

(a) The drones selected in the sample are such that every drone manufactured by the manufacturer has an equal chance of being selected.

(b) Unbiased estimate of population mean $= \frac{1171}{100} = 11.71$

Unbiased estimate of population variance, s^2

$$= \frac{1}{99} \left[14023 - \frac{1171^2}{100} \right]$$

$$= \frac{3451}{1100}$$

$$= 3.1373$$

$$= 3.14 \text{ (to 3 s.f.)}$$

(c) $H_0 : \mu = 12$

$H_1 : \mu \neq 12$, where μ is the population mean AFT of a manufactured drone in minutes.

Under H_0 ,

Since $n = 100$ is large,

$$\bar{X} \sim N\left(12, \frac{3.1373}{100}\right) \text{ approximately by CLT}$$

$$Z = \frac{\bar{X} - 12}{\sqrt{s^2/n}} \sim N(0, 1)$$

At 5% level of significance, we reject H_0 if $p\text{-value} \leq 0.05$.

$$p\text{-value} = 0.102$$

Since $p\text{-value} > 0.05$, we do not reject H_0 and conclude that there is insufficient evidence at the 5% level of significance to claim that is insufficient evidence to say that the mean AFT of the drones is not equal to 12 minutes.

(d) $H_0 : \mu = 12$
 $H_1 : \mu > 12$

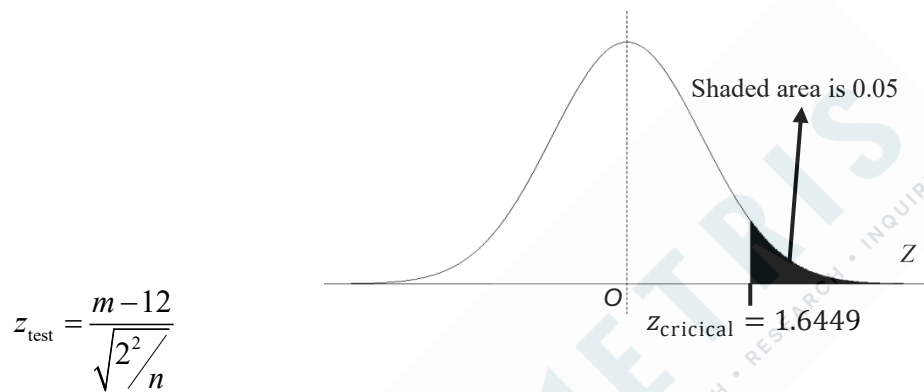
Under H_0 ,

$$\bar{X} \sim N(12, 2^2)$$

$$\bar{X} \sim N\left(12, \frac{2^2}{n}\right) \text{ approximately by CLT}$$

$$Z = \frac{\bar{X} - 12}{\sqrt{\sigma^2/n}} \sim N(0, 1)$$

At 5% level of significance, we reject H_0 if $p\text{-value} \leq 0.05$.



$$z_{\text{test}} = \frac{m - 12}{\sqrt{2^2/n}}$$

To accept that the AFT of the toy drone has in fact increased, we reject H_0 .

$$\frac{m - 12}{\sqrt{2^2/n}} \geq 1.6449$$

$$m \geq \frac{3.2897}{\sqrt{n}} + 12$$

$\therefore$ critical value is $\frac{3.29}{\sqrt{n}} + 12$ (to 3.s.f.).

12

- (a) Unbiased estimate of the population mean

$$\begin{aligned} &= \frac{-81}{120} + 27 \\ &= 26.325 \end{aligned}$$

Unbiased estimate for the population variance, s^2

$$\begin{aligned} &= \frac{1}{119} \left(2070.8 - \frac{(-81)^2}{120} \right) \\ &= 16.94222689 \\ &= 16.9 \quad (\text{to 3 s.f.}) \end{aligned}$$

- (b) $H_0 : \mu = 27$

$H_1 : \mu < 27$, where μ is the population mean mass of a batch of polypropylene per 100 grams

Under H_0 ,

Since $n = 120$ is large,

$$\begin{aligned} \bar{X} &\sim N\left(27, \frac{16.94222689}{120}\right) \text{ approximately by CLT} \\ Z &= \frac{\bar{X} - 27}{\sqrt{\frac{s^2}{n}}} \sim N(0,1) \text{ approximately.} \end{aligned}$$

At 5% level of significance, we reject H_0 if $p\text{-value} \leq 0.05$.

$$p\text{-value} = 0.0362$$

Since $p\text{-value} \leq 0.05$, we reject H_0 and conclude that there is sufficient evidence at the 5% level of significance to claim that is insufficient evidence to say that the population mean mass of polypropylene per 100g that were broken down is less than 27g.

- (c) It is the probability of wrongly concluding that the population mean mass of polypropylene per 100g that was broken down by Fungus A was less than 27g when the population mean mass is, in fact, 27 g, is 0.05.
- (d) Since the sample size of 120 is large enough, University M can use Central Limit Theorem to approximate the distribution of the sample mean mass of polypropylene per 100g that were broken down to be a normal distribution, hence University M does not need to know the distribution of the mass of polypropylene per 100g that were broken down.

(e) $H_0 : \mu = 27$

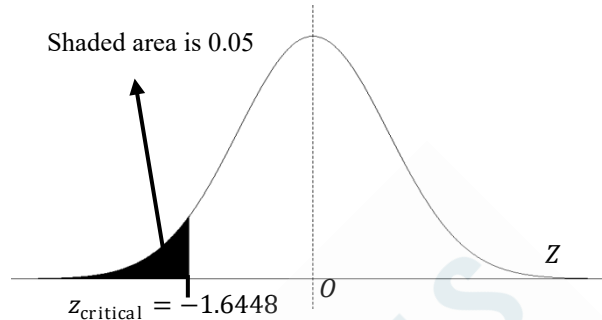
$H_1 : \mu < 27$, where μ is the population mean mass of a batch of polypropylene per 100 grams

Under H_0 ,

Since $n = 50$ is large, $\bar{Y} \sim N\left(27, \frac{4^2}{50}\right)$ approximately by CLT

$$Z = \frac{\bar{Y} - 27}{\sqrt{\sigma^2/n}} \sim N(0,1)$$

At 5% level of significance,



$$z_{\text{test}} = \frac{\bar{y} - 27}{\sqrt{4^2/50}}$$

Since null hypothesis is not rejected,

$$\frac{\bar{y} - 27}{\sqrt{4^2/50}} > -1.64485$$

$$\bar{y} - 27 > -0.9304697224$$

$$\bar{y} > 26.06953028$$

$$\therefore \{\bar{y} \in \mathbb{R} : 26.1 \leq \bar{y} \leq 100\} \text{ (to 3 s.f.)}$$

13

$$\bar{w} = \frac{5.42}{100} + 10 = 10.0542$$

$$s^2 = \frac{1}{100-1} \left[k - \frac{5.42^2}{100} \right] = \frac{1}{99} (k - 0.293764)$$

$$H_0: \mu = 10$$

$$H_1: \mu \neq 10, \text{ where } \mu \text{ is the population mean of } W$$

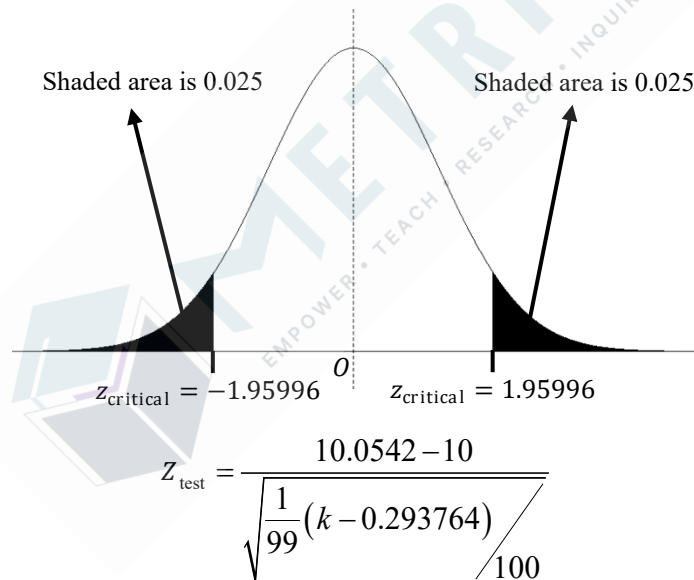
Under H_0 ,

Since $n = 100$ is large,

$$\bar{W} \sim N \left(10, \frac{\frac{1}{99} (k - 0.293764)}{100} \right) \text{ approximately by CLT}$$

$$Z = \frac{\bar{W} - 10}{\sqrt{s^2/n}} \sim N(0, 1)$$

At 5% level of significance,



Since H_0 is not rejected,

$$-1.95996 < \frac{10.0542 - 10}{\sqrt{\frac{1}{99} (k - 0.293764)} / 100} < 1.95996$$

$$-1.95996 < \frac{10.0542 - 10}{\sqrt{\frac{1}{99} (k - 0.293764)} / 100} \quad \text{and} \quad \frac{10.0542 - 10}{\sqrt{\frac{1}{99} (k - 0.293764)} / 100} < 1.95996$$

$$-1.95996 \sqrt{\frac{1}{99} (k - 0.293764)} / 100 < 0.0542 \quad \text{and} \quad \frac{10.0542 - 10}{\sqrt{\frac{1}{99} (k - 0.293764)} / 100} < 1.95996$$

(rej.)

$$k > 7.86452$$

$$k > 7.87$$

Need to assume that W follows a normal distribution.

14

(a) Unbiased estimate of population mean,

$$\begin{aligned} &= \frac{198.5}{35} + 500 \\ &= 505.671 \\ &= 506 \text{ (to 3 s.f.)} \end{aligned}$$

Unbiased estimate of population variance, s^2

$$\begin{aligned} &= \frac{1}{35-1} \left[7188 - \frac{(198.5)^2}{35} \right] \\ &= 178.3006 \\ &= 178 \text{ (to 3 s.f.)} \end{aligned}$$

(b) $H_0 : \mu = 500$

$H_1 : \mu \neq 500$ where μ is the population mean weight of a packet of scallops

Under H_0 ,

Since $n = 35$ is large,

$$\begin{aligned} \bar{X} &\sim N\left(500, \frac{178.3006}{35}\right) \text{ approximately by CLT} \\ Z &= \frac{\bar{X} - 500}{\sqrt{s^2/n}} \sim N(0, 1) \end{aligned}$$

At 2% level of significance, we reject H_0 if $p\text{-value} \leq 0.02$.

$$p\text{-value} = 0.0120$$

Since $p\text{-value} \leq 0.02$, we reject H_0 and conclude that there is sufficient evidence at the 2% level of significance to claim that the mean mass of the packets of scallops is not 500 grams.

(c) $s^2 = \frac{n}{n-1}(\text{sample variance})$
 $= \frac{40}{39}(11.7^2)$
 $= 140.4$

$$H_0 : \mu = \mu_0$$

$$H_1 : \mu < \mu_0, \text{ where } \mu \text{ is the population mean weight of a packet of scallops}$$

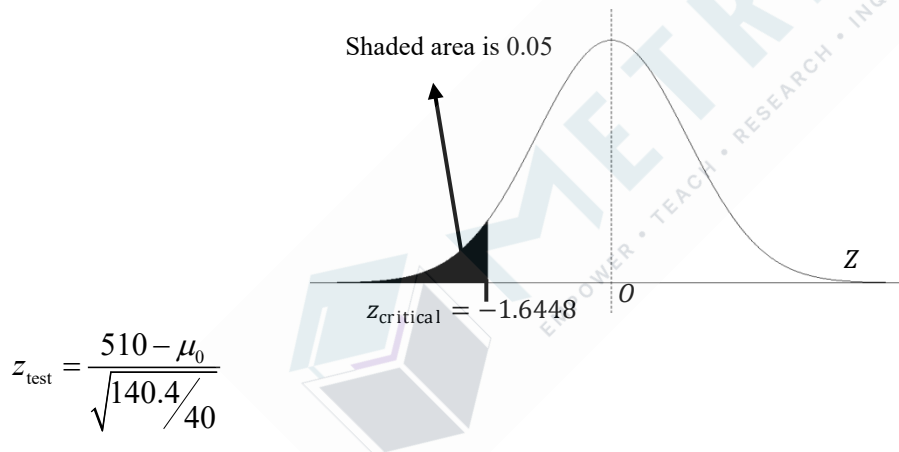
Under H_0 ,

Since $n = 40$ is large,

$$\bar{X} \sim N\left(\mu_0, \frac{140.4}{40}\right) \text{ approximately by CLT}$$

$$Z = \frac{\bar{X} - \mu_0}{\sqrt{s^2/n}} \sim N(0,1)$$

At 5% level of significance,



Since there is no reason to reject the company's claim, we do not reject H_0 ,

$$\frac{510 - \mu_0}{\sqrt{140.4/40}} > -1.64485$$

$$\mu_0 < 513.08$$

$$\therefore 0 < \mu_0 < 513.08 \text{ (to 2 d.p.)}$$

15

(i) Unbiased estimate of the population variance s^2

$$\begin{aligned} &= \frac{n}{n-1} \times \text{sample variance} \\ &= \frac{30}{29} \times 15^2 \\ &= \frac{6750}{29} \end{aligned}$$

Let X be the random variable denoting the Calculus score of a student.

$$H_0 : \mu_1 = 52$$

$H_1 : \mu_1 \neq 52$, where μ is the population mean Calculus score of a student.

Under H_0 ,

Since $n = 30$ is large,

$$\begin{aligned} \bar{X} &\sim N \left(52, \frac{6750}{29} \right) \text{ approximately by CLT} \\ Z &= \frac{\bar{X} - 52}{\sqrt{\frac{6750}{29}} / \sqrt{30}} \sim N(0,1) \end{aligned}$$

At 5% level of significance, we reject H_0 if $p\text{-value} \leq 0.05$.

$$p\text{-value} = 0.0312$$

Since $p\text{-value} \leq 0.05$, we reject H_0 and conclude that there is sufficient evidence at the 5% level of significance to conclude that Professor's claim about the mean score is not valid.

- (ii) Let Y be the random variable denoting the Statistics score by a student and μ be the population mean.

$$H_0 : \mu = 48$$

$$H_1 : \mu > 48, \text{ where } \mu \text{ is the population mean Statistics score of a student.}$$

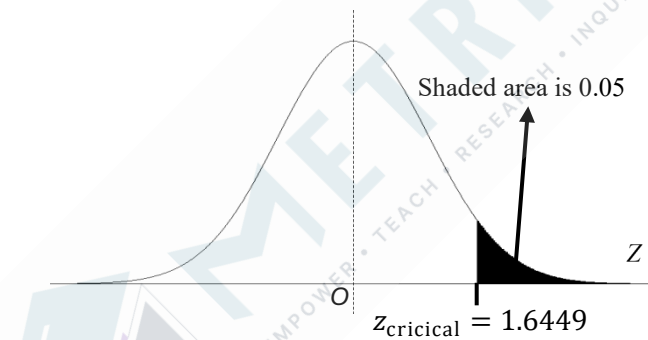
Under H_0 ,

Since $n = 30$ is large,

$$\bar{Y} \sim N\left(48, \frac{13^2}{30}\right) \text{ approximately by CLT}$$

$$Z = \frac{\bar{Y} - 48}{\sqrt{\sigma^2/n}} \sim N(0, 1)$$

At 5% level of significance, we reject H_0 if $p\text{-value} \leq 0.05$.



$$z_{\text{test}} = \frac{k - 48}{\sqrt{13^2/30}}$$

To conclude that the professor has indeed understated the mean score, we reject H_0 ,

$$\frac{k - 48}{\sqrt{13^2/30}} \geq 1.6449$$

$$k \geq 51.904$$

$$k \geq 52.0 \text{ (to 3 s.f.)}$$

- (iii) Yes. Since it is not known whether the Calculus score and Statistics score by a randomly selected student is normally distributed, it is important that the sample size is at least 30 in order to use Central Limit Theorem so that the mean Calculus score and the mean Statistics score by students are approximately normal.