The University of Sydney

MATH1005 Statistics

maths.usyd.edu.au/u/UG/WS/WS1005/

Summer/Winter/Semester2 **Tutorial Solutions 11** 2015

This tutorial explores the 1 sample Z test.

1 Sample Z Test

Context	n observations from a population with unknown mean μ and known variance σ^2
Hypothesis	$H_0 : \mu = \mu_0$
Test Statistic	$\tau = \dfrac{\bar{X} - \mu_0}{\frac{\sigma}{\sqrt{n}}} \overset{H_0}{\sim} N(0, 1)$

1. In what contexts can you use the Z test? Is this likely?

Solution

Context for the Z test: When the population variance σ^2 is known. This is not very common. Hence, we often need the t Test (Tute 12).

2. The breaking strengths of ropes produced by a manufacturer have mean 1800N and standard deviation 100N. By a new technique in the manufacturing process, it is claimed that the breaking strength can be increased without changing the variability. To test this claim a sample of 50 ropes is tested and it is found that the mean breaking strength is 1850N. Can you support the manufacturer's claim? Justify your decision using a statistical argument.

Let $X_i\ \ i = 1, \ldots, 50$ be the breaking strengths and let μ be the population mean of the ropes produced by the new technique.

(a) Hypothesis: Explain why a one-sided test of $H_0 : \mu = 1800\ \ vs\ \ H_1 : \mu > 1800$ is suitable to test the manufacturer's claim.

(b) Assumptions: Are the assumptions for a Z test valid here?

(c) Test statistic: What is the test statistic and its distribution under H_0?

(d) P-value: Calculate the p-value using the Normal tables and R.

```
> 1-pnorm(1850,1800,100/sqrt(50))
```

(e) Conclusion: Draw your conclusion based on the p-value.

Solution

Let $X_i\ \ i = 1, \ldots, 50$ be the breaking strengths and let μ be the population mean of the ropes produced by the new technique. We have a sample with $n = 50$ and $\bar{x} = 1850$.

$\boxed{\text{H}}$ $H_0 : \mu = 1800\ \ vs\ \ H_1 : \mu > 1800$
This one-sided test is suitable as the manufacturer claims that rope strength has *increased.*

$\boxed{\text{A}}$ $\sigma^2 = 100^2$ is known ('new technique ... without changing the variability').

$\boxed{\text{T}}$ $\tau = \bar{X} \overset{H_0}{\sim} N(1800, 100^2/50)$.
Observed value: $\bar{x} = 1850$.

OR in standardised form:

$$\tau = \frac{\bar{X} - 1800}{\frac{100}{\sqrt{50}}} \overset{H_0}{\sim} N(0,1) = Z.$$

Observed value: $\tau_0 = \dfrac{1850 - 1800}{\frac{100}{\sqrt{50}}} = 3.535534$

$\boxed{\text{P}}$

Using Normal Tables:
P-value $= P(Z > 3.535534) = 1 - P(Z < 3.54) = 0.0002.$

Using R:

```
> 1-pnorm(1850,1800,100/sqrt(50))
[1] 0.000203476
OR
> 1-pnorm(3.54)
[1] 0.0002000635
```

$\boxed{\text{C}}$ Since p-value $<< 0.05$, we strongly reject H_0.
There appears to be strong evidence that the rope strength has increased.

3. The following data represents marks for a group of 18 students, for a test out of 10.

```
1 2 3 3 4 5 5 5 6 6 6 7 7 8 9 9 9 10
```

(a) By hand, show that the mean is $\bar{x} = 5.83$ and the standard devation is $s = 2.60$ (2dp). Confirm using R.

```
> x=c(1,2,3,3,4,5,5,5,6,6,6,7,7,8,9,9,9,10)
> mean(x)
> sd(x)
```

(b) Produce a boxplot of the data. Comment.

(c) Perform a Z test for $H_0 : \mu = 5$ vs $H_1 : \mu \neq 5$, showing all your steps.

Solution

Let the student marks be $\{x_i\}$.

(a) $\displaystyle\sum_{i=1}^{18} x_i = 1 + 2 + \ldots 10 = 105.$

$\displaystyle\sum_{i=1}^{18} x_i^2 = 1^2 + 2^2 + \ldots 10^2 = 727.$

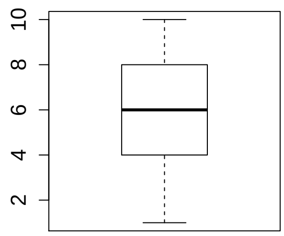
Mean: $\bar{x} = 105/18 \doteq 5.83.$

Sd: $s = \sqrt{\dfrac{1}{17}(727 - \dfrac{1}{18}(105)^2)} \doteq 2.60.$

Confirm using R:

```
> x=c(1,2,3,3,4,5,5,5,6,6,6,7,7,8,9,9,9,10)
> mean(x)
[1] 5.833333
> sd(x)
[1] 2.595245
```

(b) Comment: Tne boxplot looks symmetric, which would be consistent with a Normal population.



(c) Let $X_i \;\; i = 1, \ldots, 18$ be the student marks and let μ be the population mean of the marks, where $\bar{x} = 5.83$ and $s = 2.60$.

$\boxed{\text{H}}$ $H_0 : \mu = 5 \;\; vs \;\; H_1 : \mu \neq 5$

$\boxed{\text{A}}$ Assume that $\sigma^2 = s^2 = 2.595245^2 \doteq 6.74$.

$\boxed{\text{T}}$ $\tau = \dfrac{\bar{X} - 5}{\frac{2.595245}{\sqrt{18}}} \overset{H_0}{\sim} N(0, 1) = Z.$

Observed value: $\tau_0 = \dfrac{5.83333 - 5}{\frac{2.595245}{\sqrt{18}}} = 1.362307$

$\boxed{\text{P}}$

Using Normal Tables:
P-value $= 2P(Z > 1.36) = 2(1 - P(Z < 1.36)) = 2(1 - 0.9131) = 0.1738$.

Using R:

```
> 2*(1-pnorm(0.3210872))
[1] 0.1738299
```

$\boxed{\text{C}}$ Since p-value > 0.05, we retain H_0.
Data is consistent with $\mu = 5$.

4. A standard manufacturing process produces items whose lengths in mm are normally distributed with mean $\mu = 10$ and standard deviation $\sigma = 0.051$. A new, cheaper manufacturing process is being tested, but a manager is worried that the new process makes the items too short. A sample of 100 items produced by the new process has average length $\bar{x} = 9.9921$. Assume the variability in lengths of the new process is the same as the old. Perform a statistical test to address the manager's concerns.

Solution

Let $X_i \;\; i = 1, \ldots, 18$ be the item lengths and let μ be the population mean of the lengths and $\sigma = 0.051$. The sample of $n = 100$ gives $\bar{x} = 9.9921$.

$\boxed{\text{H}}$ $H_0 : \mu = 10 \;\; vs \;\; H_1 : \mu < 10$

$\boxed{\text{A}}$ We assume that $\sigma = 0.051$ (as 'variability of lengths of the new process is the same as the old').

$\boxed{\text{T}}$ $\tau = \dfrac{\bar{X} - 10}{\frac{0.051}{\sqrt{100}}} \overset{H_0}{\sim} N(0, 1) = Z.$

Observed value: $\tau_0 = \dfrac{9.9921 - 10}{\frac{0.051}{\sqrt{100}}} = \dfrac{9.9921 - 10}{0.0051} = -1.54902$

$\boxed{\text{P}}$

Using Normal Tables:
P-value $= P(Z < -1.55) = 1 - P(Z < 1.55) = 1 - 0.9394 = 0.0606$.

Using R:
```
> pnorm(-1.55)
[1] 0.06057076
```

$\boxed{\text{C}}$ Since p-value > 0.05 (just), we retain H_0.
There is not enough evidence to suggest that the manager's claim is unfounded. However, there is a slight indication that further research might be worthwhile.

5. A random sample of size 64 from a $N(\mu, 16)$ population yields a sample mean of 19.1. Test $H_0\colon \mu = 20$ against $H_1\colon \mu < 20$. Write all your steps.

Solution

Let $X_i \;\; i = 1, \ldots, 64$ be the sample and let μ be the population mean with $\sigma^2 = 16$. The sample of $n = 64$ gives $\bar{x} = 19.1$.

$\boxed{\text{H}}$ $H_0 : \mu = 20 \;\; vs \;\; H_1 : \mu < 20$

$\boxed{\text{A}}$ We are given $\sigma = 4$.

$\boxed{\text{T}}$ $\tau = \dfrac{\bar{X} - 20}{\frac{4}{\sqrt{64}}} = \dfrac{\bar{X} - 20}{0.5} \overset{H_0}{\sim} N(0, 1) = Z.$

Observed value: $\tau_0 = \dfrac{19.1 - 20}{0.5} = -1.8$

$\boxed{\text{P}}$

Using Normal Tables:
P-value $= P(Z < -1.8) = 1 - P(Z < 1.8) = 1 - 0.9641 = 0.0359$.

Using R:
```
> pnorm(-1.8)
[1] 0.03593032
```

$\boxed{\text{C}}$ Since p-value < 0.05, we reject H_0 in favour of H_1.

6. A sample of size 22, from a normal population with known variance of 2.25, yields a total of 250. Find an estimate of the population mean μ. Test the hypothesis that the population mean μ is 10 against the alternative that the mean is greater than 10. Write all your steps.

Solution

Let $X_i \;\; i = 1, \ldots, 22$ be the sample and let μ be the population mean with $\sigma^2 = 2.25$. The sample of $n = 22$ gives $\displaystyle\sum_{i=1}^{22} x_i = 250$.

$\boxed{\text{H}}$ $H_0 : \mu = 10 \;\; vs \;\; H_1 : \mu > 10$

$\boxed{\text{A}}$ We are given $\sigma^2 = 2.25$.

$\boxed{\text{T}}$ $\tau = \dfrac{\bar{X} - 10}{\frac{\sqrt{2.25}}{\sqrt{22}}} \overset{H_0}{\sim} N(0, 1) = Z.$

Observed value: $\tau_0 = \dfrac{\frac{250}{22} - 10}{\frac{\sqrt{2.25}}{\sqrt{22}}} = 3.501168$

$\boxed{\text{P}}$

Using Normal Tables:
P-value $= P(Z > 3.50) = 1 - P(Z < 3.50) = 1 - 0.9998 = 0.0002$.

Using R:
```
> 1-pnorm(3.5)
[1] 0.0002326291
```

$\boxed{\text{C}}$ Since p-value $<< 0.05$, we reject H_0 in favour of H_1.

7. Suppose that $X_1, X_2, ..., X_{49}$ are a $N(\mu, 10^2)$ random sample, for some unknown μ. Test the hypothesis $H_0 : \mu = 50$ against the alternative $H_1 : \mu < 50$ based on an observed value $\bar{x} = 48.8$, being sure to give a sensible interpretation of your p-value. Write down all your steps.

Solution

Let $X_i \;\; i = 1, \ldots, 49$ be the sample and let μ be the population mean with $\sigma^2 = 10^2$. The sample of $n = 49$ gives $\bar{x} = 48.8$.

$\boxed{\text{H}}$ $H_0 : \mu = 50 \;\; vs \;\; H_1 : \mu < 50$

$\boxed{\text{A}}$ We are given $\sigma^2 = 10^2$.

$\boxed{\text{T}}$ $\tau = \dfrac{\bar{X} - 50}{\frac{10}{7}} \overset{H_0}{\sim} N(0, 1) = Z.$

Observed value: $\tau_0 = \dfrac{48.8 - 50}{10/7} = -0.84$

$\boxed{\text{P}}$

Using Normal Tables:
P-value $= P(Z < -0.84) = 1 - P(Z < 0.84) = 1 - 0.7995 = 0.2005$.

Using R:
```
> pnorm(-0.84)
[1] 0.2004542
```

$\boxed{\text{C}}$ Since p-value > 0.05, we retain H_0.

8. In accordance with the standards established for a reading comprehension test, year 8 students should average a score of 84.3 on a standard test with a standard deviation of 8.6. If 36 randomly selected eighth graders (from a school district in a low socio-economic area) averaged only 80.8, is there strong evidence that the average in that area is below standard, if it is assumed that the standard deviation for this area is 8.6, in accordance with the standard for eighth graders?

Solution

Let $X_i \;\; i = 1, \ldots, 36$ be the sample and let μ be the population mean with $\sigma^2 = 8.6^2$. The sample of $n = 49$ gives $\bar{x} = 80.8$.

$\boxed{\text{H}}$ $H_0 : \mu = 84.3 \;\; vs \;\; H_1 : \mu < 84.3$

$\boxed{\text{A}}$ We are given $\sigma = 8.6$.

$\boxed{\text{T}}$ $\tau = \dfrac{\bar{X} - 84.3}{\frac{8.6}{6}} \overset{H_0}{\sim} N(0,1) = Z.$

Observed value: $\tau_0 = \dfrac{80.8 - 84.3}{8.6/6} = -2.44186$

$\boxed{\text{P}}$

Using Normal Tables:
P-value $= P(Z < -2.44) = 1 - P(Z < 2.44) = 1 - 0.9927 = 0.0073.$

Using R:

```
> pnorm(-2.44)
[1] 0.007343631
```

$\boxed{\text{C}}$ Since p-value $<< 0.05$, we reject H_0 in favour of H_1.
It seems that reading in the low socio-economic area is below average.