Hypothesis Testing

Illustrated through inference on one population mean or proportion

Part 1. Inference on one population mean

Scenario 1. When the population is normal, and the population variance is known

$$\text{Data}: X_1, X_2, \ldots, X_n \overset{i.i.d.}{\sim} N(\mu, \sigma^2)$$

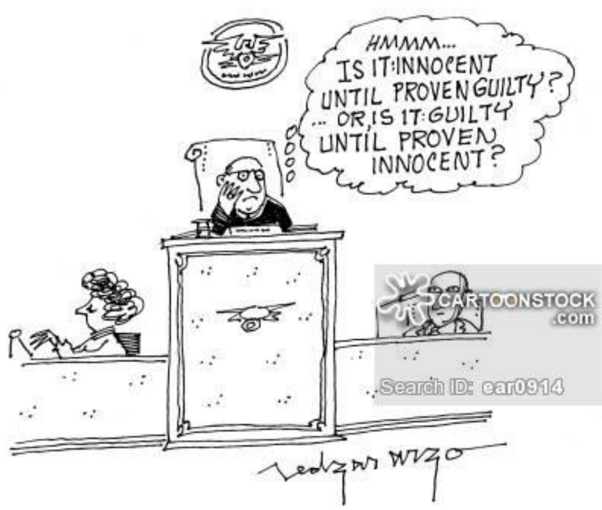
Hypothesis test, for instance: $\begin{cases} H_0: \mu = \mu_0 \\ H_a: \mu > \mu_0 \end{cases} \Leftrightarrow \begin{cases} H_0: \mu \le \mu_0 \\ H_a: \mu > \mu_0 \end{cases}$

Example:

$H_0: \mu \le 5'7''$ (**null hypothesis**): This is the 'original belief'

$H_a: \mu > 5'7''$ (**alternative hypothesis**) : This is usually your hypothesis (i.e. what you believe is true) if you are conducting the test – and in general, should be supported by your data.

The statistical hypothesis test is very similar to a law suit:



e.g) The famous O.J. Simpson trial

H_0: OJ is innocent ('innocent unless proven guilty)

H_a: OJ is guilty ('supported by the data: the evidence)

		The truth	
		H_0: OJ innocent	H_a OJ guilty
Jury's Decision	H_0	Right decision	Type II error
	H_a	Type I error	Right decision

The significance level and three types of hypotheses.

P(Type I error) = α ← significance level of a test (*Type I error rate)

1. $H_0: \mu = \mu_0$ ⇔ $H_0: \mu \le \mu_0$

   $H_a: \mu > \mu_0$ $\quad$ $H_a: \mu > \mu_0$

2. $H_0: \mu = \mu_0$ ⇔ $H_0: \mu \ge \mu_0$

   $H_a: \mu < \mu_0$ $\quad$ $H_a: \mu < \mu_0$

3. $H_0: \mu = \mu_0$

   $H_a: \mu \ne \mu_0$

<u>Now we derive the hypothesis test for the first pair of hypotheses.</u>

$H_0: \mu = \mu_0$

$H_a: \mu > \mu_0$

Data : $X_1, X_2, \ldots, X_n \overset{iid}{\sim} N(\mu, \sigma^2), \sigma^2$ is known and the given a significance level α (say, 0.05). Let's derive the test. (That is, derive the decision rule)

Two approaches (*equivalent) to derive the tests:

- Likelihood Ratio Test
- Pivotal Quantity Method

***Now we will first demonstrate the **Pivotal Quantity** Method.

1. We have already derived the PQ when we derived the C.I. for μ

$$Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim N(0,1) \text{ is our P.Q.}$$

2. The test statistic is the PQ with the value of the parameter of interest under the null hypothesis (H_0) inserted:

$$Z_0 = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \overset{H_0}{\sim} N(0,1) \text{ is our test statistic.}$$

That is, given $H_0 : \mu = \mu_0$ in true $\Rightarrow Z_0 = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \sim N(0,1)$

3. * Derive the decision threshold for your test based on the Type I error rate the significance level α

For the pair of hypotheses:

$H_0 : \mu = \mu_0$ **versus** $H_a : \mu > \mu_0$

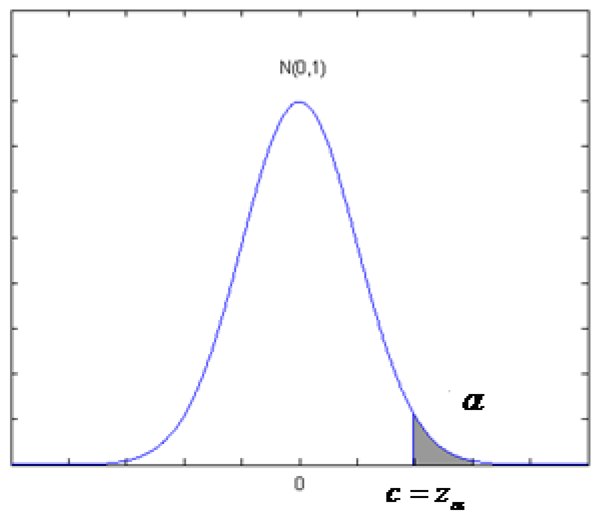
It is intuitive that one should reject the null hypothesis, in support of the alternative hypothesis, when the sample mean is larger than μ_0. Equivalently, this means when the test statistic Z_0 is larger than certain positive value c - the question is what is the exact value of c -- and that can be determined based on the significance level α—that is, how much Type I error we would allow ourselves to commit.

Setting:

P(Type I error) = P(reject $H_0 \mid H_0$) =

$P(Z_0 \geq c \mid H_0 : \mu = \mu_0) = \alpha$

We will see immediately that $c = z_\alpha$ from the pdf plot below.



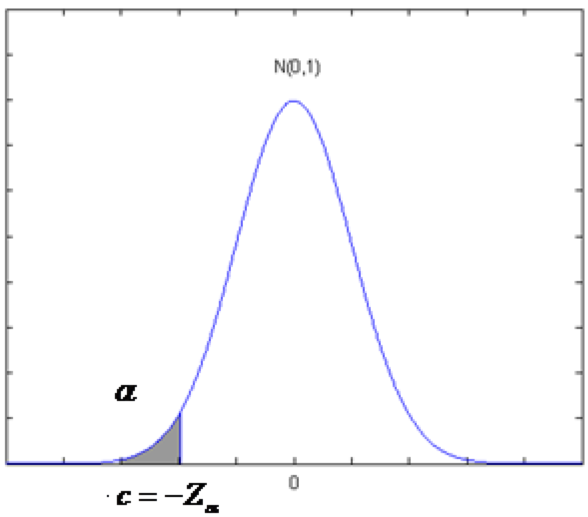
∴ At the significance level α, we will reject H_0 in favor of H_a if $Z_0 \geq Z_\alpha$

Other Hypotheses

$H_0: \mu = \mu_0$
$H_a: \mu < \mu_0$ (one-sided test or one-tailed test)

$$\text{Test statistic}: Z_0 = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \sim N(0,1)$$

$$\alpha = P(Z_0 \le c \mid H_0 : \mu = \mu_0) \Rightarrow c = -Z_\alpha$$



$H_0: \mu = \mu_0$
$H_a: \mu \neq \mu_0$ (Two-sided or Two-tailed test)

$$\text{Test statistic}: Z_0 = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \sim N(0,1)$$

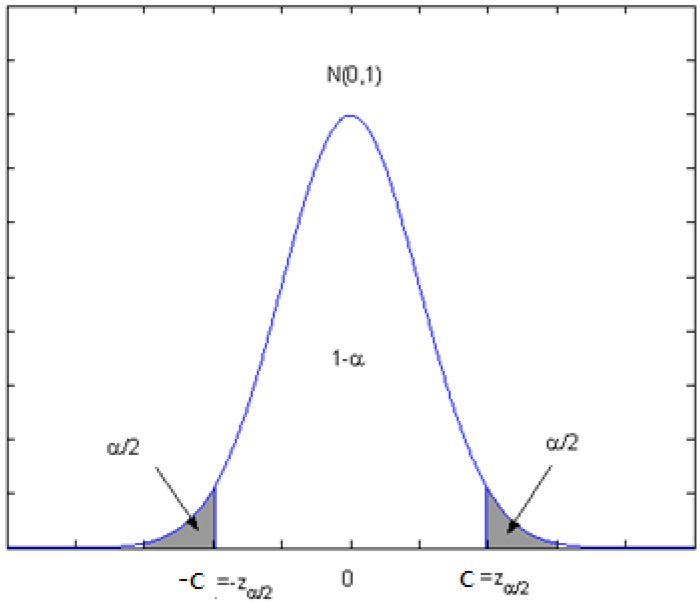
$$\alpha = P(|Z_0| \ge c \mid H_0) = P(Z_0 \ge c \mid H_0) + P(Z_0 \le -c \mid H_0)$$

$$= 2 \cdot P(Z_0 \ge c \mid H_0)$$

$$\frac{\alpha}{2} = P(Z_0 \ge c \mid H_0)$$

$$\therefore c = Z_{\alpha/2}$$

Reject H_0 if $|Z_0| \ge Z_{\alpha/2}$



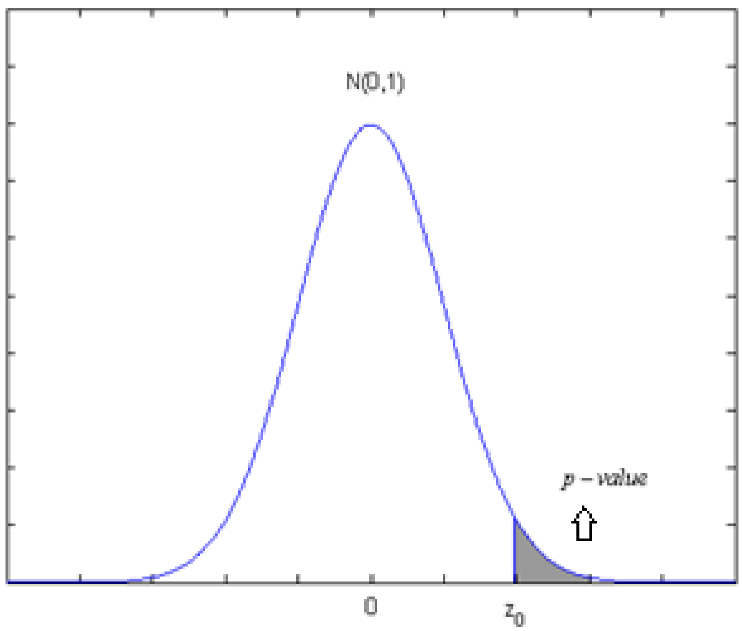
4. We have just discussed the ***"rejection region"*** approach for decision making.

There is another approach for decision making, it is ***"p-value"*** approach.

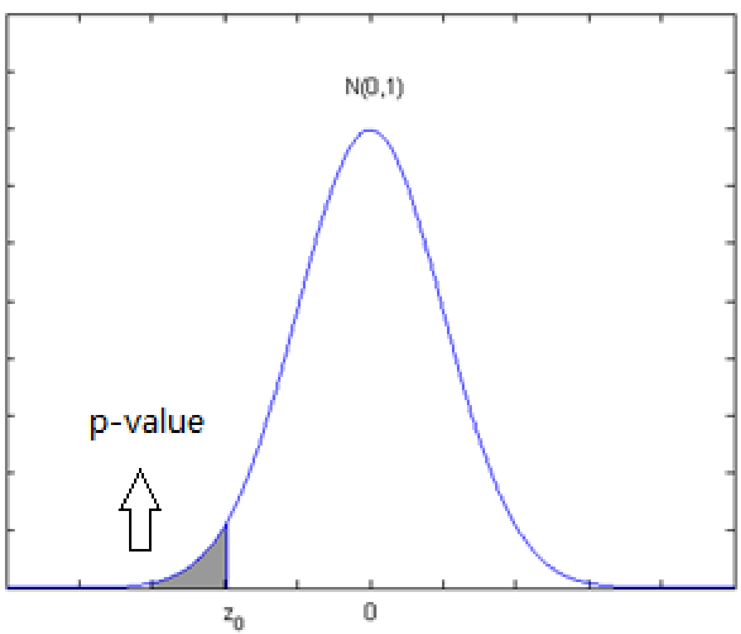
*Definition: p-value – it is the probability that we observe a test statistic value that is as extreme, or more extreme, than the one we observed, given that the null hypothesis is true.

$H_0:\mu=\mu_0$ $H_a:\mu>\mu_0$	$H_0:\mu=\mu_0$ $H_a:\mu<\mu_0$	$H_0:\mu=\mu_0$ $H_a:\mu\neq\mu_0$
Observed value of test statistic $Z_0=\dfrac{\overline{X}-\mu_0}{\sigma/\sqrt{n}}\overset{H_0}{\sim}N(0,1)$		
p-value $=P(Z_0\geq z_0 \mid H_0)$	p-value $=P(Z_0\leq z_0 \mid H_0)$	p-value $=P(\lvert Z_0\rvert\geq\lvert z_0\rvert \mid H_0)$ $=2\cdot P(Z_0\geq\lvert z_0\rvert \mid H_0)$
(1) the area under N(0,1) pdf to the right of z_0	(2) the area under N(0,1) pdf to the left of z_0	(3) twice the area to the right of $\lvert z_0\rvert$

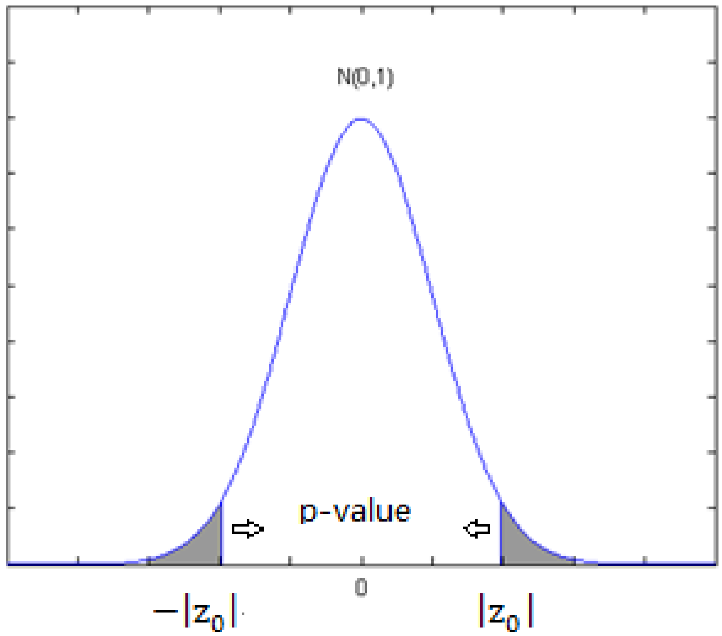
(1) $\begin{array}{l} H_0 : \mu = \mu_0 \\ H_a : \mu > \mu_0 \end{array}$



(2) $\begin{array}{l} H_0 : \mu = \mu_0 \\ H_a : \mu < \mu_0 \end{array}$

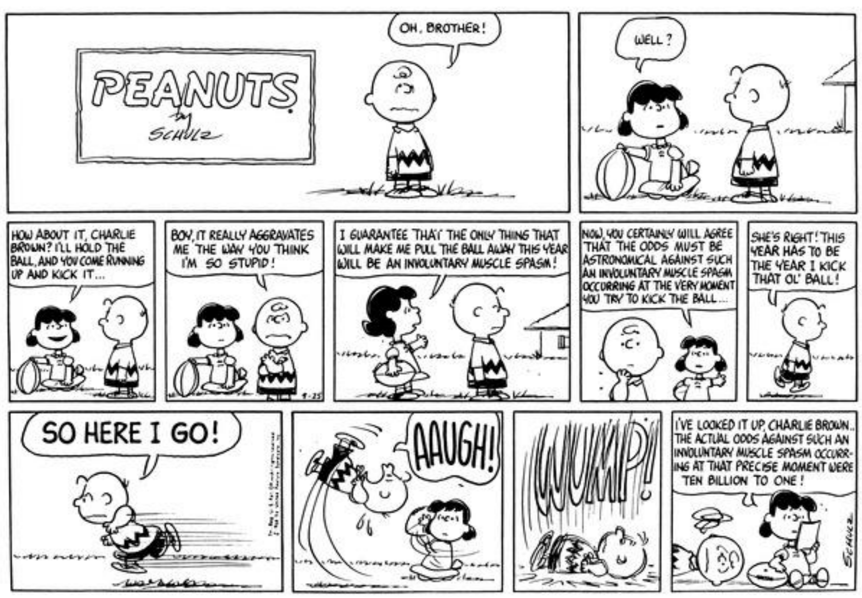


(3) $\begin{array}{l} H_0 : \mu = \mu_0 \\ H_a : \mu \neq \mu_0 \end{array}$



The way we make conclusions is the **same** for all hypotheses:

We reject H_0 in favor of H_a iff p-value < α



The experimental evidence against the null hypothesis that Lucy did not act deliberately reaches a P-value of one in ten billion. Nevertheless, Charlie Brown repeats the experiment every year.

Scenario 2. **The large sample scenario: Any population (*usually non-normal – as the exact tests should be used if the population is normal), however, the sample size is large** ***(this usually refers to: n ≥ 30)***

Theorem. The Central Limit Theorem

Let $X_1, X_2, \cdots, X_n$ be a random sample from a population with mean μ and variance σ^2

$$\frac{\overline{X} - \mu}{\sigma / \sqrt{n}} \xrightarrow{n \to \infty} N(0,1)$$

* When n is large enough $(n \ge 30)$,

$\boxed{Z = \frac{\overline{X} - \mu}{S / \sqrt{n}} \overset{\bullet}{\sim} N(0,1)}$ *(approximately) – by CLT and the Slutsky's Theorem*

Therefore the pivotal quantities **(P.Q.'s) for this scenario:**

$$Z = \frac{\overline{X} - \mu}{\sigma / \sqrt{n}} \overset{\bullet}{\sim} N(0,1) \text{ or } Z = \frac{\overline{X} - \mu}{S / \sqrt{n}} \overset{\bullet}{\sim} N(0,1)$$

Use the first P.Q. if σ is known, and the second when σ is unknown.

<u>The derivation of the hypothesis tests (rejection region and the p-value) are almost the same as the derivation of the exact Z-test discussed above.</u>

<table>
<tr><td>$H_0: \mu = \mu_0$
$H_a: \mu > \mu_0$</td><td>$H_0: \mu = \mu_0$
$H_a: \mu < \mu_0$</td><td>$H_0: \mu = \mu_0$
$H_a: \mu \neq \mu_0$</td></tr>
<tr><td colspan="3">Test Statistic $Z_0 = \dfrac{\overline{X} - \mu_0}{S/\sqrt{n}} \overset{\bullet}{\sim} N(0,1)$</td></tr>
<tr><td colspan="3">Rejection region : we reject H_0 in favor of H_a at the significance level α if</td></tr>
<tr><td>$Z_0 \geq Z_\alpha$</td><td>$Z_0 \leq -Z_\alpha$</td><td>$|Z_0| \geq Z_{\alpha/2}$</td></tr>
<tr><td>p-value
$= P(Z_0 \geq z_0 \mid H_0)$</td><td>p-value
$= P(Z_0 \leq z_0 \mid H_0)$</td><td>p-value
$= P(|Z_0| \geq |z_0| \mid H_0)$
$= 2 \cdot P(Z_0 \geq |z_0| \mid H_0)$</td></tr>
<tr><td>(1) the area under N(0,1) pdf to the right of z_0</td><td>(2) the area under N(0,1) pdf to the left of z_0</td><td>(3) twice the area to the right of $|z_0|$</td></tr>
</table>

Scenario 3. Normal Population, but the population variance is unknown

100 years ago – people use Z-test

This is OK for n large $(n \geq 30) \Rightarrow$ per the CLT (Scenario 2)

This is NOT ok if the sample size is samll.

"A <u>Student</u> of Statistics"
– pen name of **William Sealy Gosset** (June 13, 1876–October 16, 1937)

"The Student's t-test"

P.Q. $\boxed{T = \dfrac{\overline{X} - \mu}{S/\sqrt{n}} \sim t_{n-1}}$

(Exact t-distribution with n-1 degrees of freedom)

"A Student of Statistics" – pen name of **William Sealy Gosset**
(June 13, 1876–October 16, 1937)
http://en.wikipedia.org/wiki/William_Sealy_Gosset

Review: Theorem Sampling from the normal population

Let $X_1, X_2, \cdots, X_n \overset{i.i.d.}{\sim} N(\mu, \sigma^2)$, then

1) $\bar{X} \sim N(\mu, \frac{\sigma^2}{n})$

2) $W = \frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1}$

3) $\bar{X}$ and S^2 (and thus W) are independent. Thus we have:

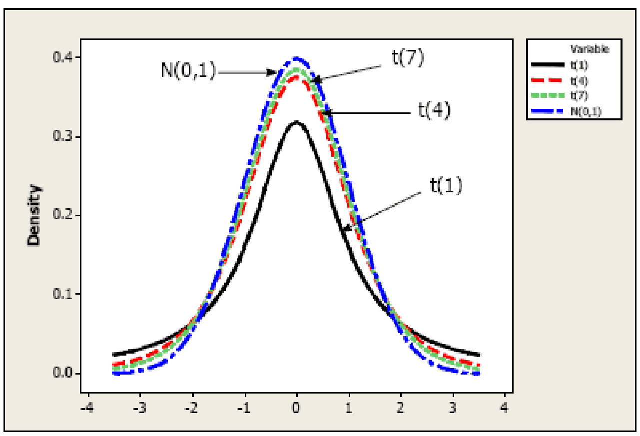
$$T = \frac{\bar{X} - \mu}{S/\sqrt{n}} \sim t_{n-1}$$

Wrong Test for a 2-sided alternative hypothesis

Reject H_0 if $|z_0| \geq Z_{\alpha/2}$

Right Test for a 2-sided alternative hypothesis

Reject H_0 if $|t_0| \geq t_{n-1,\alpha/2}$



(Because t distribution has heavier tails than normal distribution.)

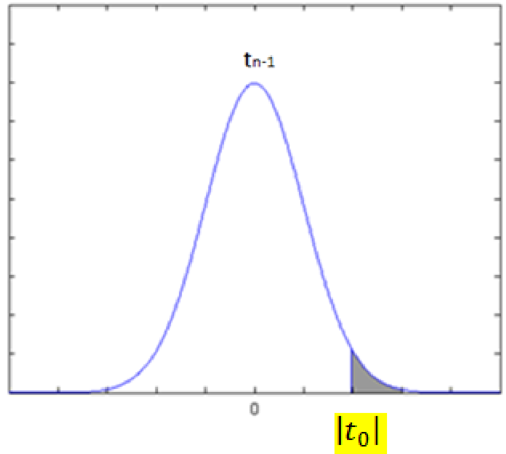
Right Test

* Test Statistic $\begin{array}{l} H_0 : \mu = \mu_0 \\ H_a : \mu \neq \mu_0 \end{array}$

$$T_0 = \frac{\bar{X} - \mu_0}{S/\sqrt{n}} \overset{H_0}{\sim} t_{n-1}$$

* Reject region : Reject H_0 at α if **the observed test statistic value** $|t_0| \geq t_{n-1,\alpha/2}$

* p-value



p-value = shaded area * 2

Further Review:

5. Definition : t-distribution

$$T = \frac{Z}{\sqrt{W/k}} \sim t_k$$

$Z \sim N(0,1)$

$W \sim \chi_k^2$ (chi-square distribution with k degrees of freedom)

$Z \,\&\, W$ are independent.

6. Def 1 : chi-square distribution : from the definition of the gamma distribution: gamma(α = k/2, β = 2)

MGF: $M(t) = \left(\frac{1}{1-2t}\right)^{k/2}$

mean & varaince: $E(W) = k; Var(W) = 2k$

Def 2 : chi-square distribution : Let

$$Z_1, Z_2, \cdots, Z_k \overset{i.i.d.}{\sim} N(0,1),$$

then $W = \sum_{i=1}^{k} Z_i^2 \sim \chi_k^2$

Example Jerry is planning to purchase a sports good store. He calculated that in order to cover basic expenses, the average daily sales must be at least $525.

Scenario A. He checked the daily sales of 36 randomly selected business days, and found the average daily sales to be $565 with a standard deviation of $150.

Scenario B. Now suppose he is only allowed to sample 9 days. And the 9 days sales are $510, 537, 548, 592, 503, 490, 601, 499, 640.

For A and B, please determine whether Jerry can conclude the daily sales to be at least $525 at the significance level of $\alpha = 0.05$. What is the p-value for each scenario?

Solution Scenario A large sample (⑤) n=36, $\bar{x} = 565, s = 150$

$$H_0 : \mu = 525 \text{ versus } H_a : \mu \succ 525$$

*** First perform the Shapiro-Wilk test to check for normality. If normal, use the exact T-test. If not normal, use the large sample Z-test. In the following, we assume the population is found not normal.

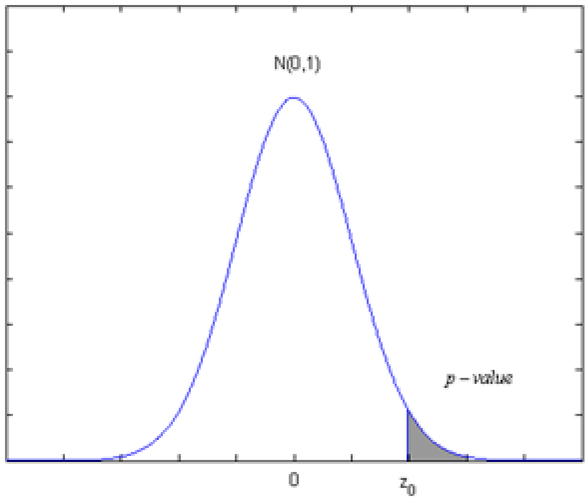
Test statistic $z_0 = \dfrac{\bar{x} - \mu_0}{s/\sqrt{n}} = \dfrac{565 - 525}{150/\sqrt{36}} = 1.6$

At the significance level $\alpha = 0.05$, we will reject H_0 if

$z_0 \geq Z_{0.05} = 1.645$

$\therefore$ We can not reject H_0

p-value



p-value = 0.0548

Alternatively, if you can show the population is normal using the Shapiro-Wilk test, it is better that you perform the exact t-test.

Solution Scenario B small sample $\Rightarrow$ Shapiro-Wilk test

$\Rightarrow$ If the population is normal, t-test is suitable.

(*If the population is not normal, and the sample size is small, we shall use the non-parametric test such as Wilcoxon Signed Rank test.)

In the following, we assume the population is found normal.

$$\bar{x} = 546.67, s = 53.09, n = 9$$

$$H_0 : \mu = 525 \text{ versus } H_a : \mu \succ 525$$

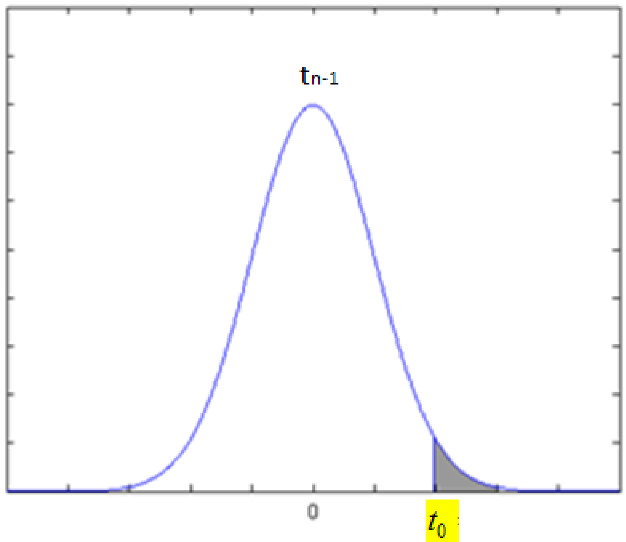
Test statistic $t_0 = \dfrac{\bar{x} - \mu_0}{s/\sqrt{n}} = \dfrac{546.67 - 525}{53.09/\sqrt{9}} = 1.22$

At the significance level $\alpha = 0.05$, we will reject H_0 if

$t_0 \geq T_{8,0.05} = 1.86$

$\therefore$ We can not reject H_0

p-value



What's the p-value when $t_0 = 1.22$?

Part 2. Inference on one population proportion

Sampling from the Bernoulli population

Toss a coin, and obtain the result as following: Head (H), H, Tail (T), T, T, H, …

Let $X_i = \begin{cases} 1, & \text{if the } i^{th} \text{ toss is head.} \\ 0, & \text{if the } i^{th} \text{ toss is tail.} \end{cases}$

A proportion of p is head, in the population.

$X_i \sim Bernoulli(p) \;\Rightarrow\; P(X_i = x_i) = p^{x_i}(1-p)^{1-x_i}\,,\quad x_i = 0,\,1$

(*Binomial distribution with n = 1)

Inference on p – the population proportion:

① Point estimator:

$\hat{p} = \dfrac{X}{n} = \dfrac{\sum_{i=1}^{n} X_i}{n}$, $\hat{p}$ is the sample proportion and also, the sample mean

② Large sample inference on p:

$$\hat{p} \sim N\left(p, \frac{p(1-p)}{n}\right)$$

③ Large sample (the original) pivotal quantity for the inference on p.

$$\boxed{Z = \frac{\hat{p}-p}{\sqrt{\dfrac{p(1-p)}{n}}} \sim N(0,1)}$$

Alternative P.Q.

$$\boxed{Z^* = \frac{\hat{p} - p}{\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}} \dot{\sim} N(0,1)}$$

④ Derive the 100(1-α)% large sample C.I. for p (using the alternative P.Q.):

If n is large, by the CLT and the Slutsky's Theorem, we have the following pivotal quantity for the inference on p:

$$\boxed{Z^* = \frac{\hat{p} - p}{\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}} \dot{\sim} N(0,1)}$$

$$P(-z_{\alpha/2} \le Z^* \le z_{\alpha/2}) = 1-\alpha$$

$$P\left(-z_{\alpha/2} \le \frac{\hat{p} - p}{\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}} \le z_{\alpha/2}\right) = 1-\alpha$$

$$P\left(-\hat{p} - z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \le -p \le -\hat{p} + z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\right) = 1-\alpha$$

$$P\left(\hat{p} - z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \le p \le \hat{p} + z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\right) = 1-\alpha$$

Hence, the 100(1-α)% large sample CI for p is

$$\hat{p} \pm z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\ .$$

⑤ Large Sample Test (via the P.Q. method):

$$\begin{cases} H_0 : p = p_0 \\ H_a : p > p_0 \end{cases} \qquad \begin{cases} H_0 : p = p_0 \\ H_a : p < p_0 \end{cases}$$

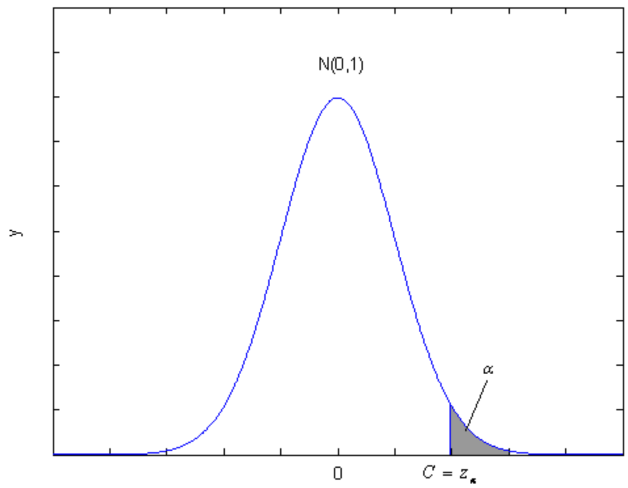
$$\begin{cases} H_0 : p = p_0 \\ H_a : p \ne p_0 \end{cases}$$

Test statistic (large sample – using the original P.Q.):

$$\boxed{Z_0 = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} \overset{H_0}{\sim} N(0,1)}$$

At the significance level α, we reject H_0 in favor of $H_a : p > p_0$ if $z_0 \geq z_\alpha$.

$$\alpha = P(reject\, H_0 \mid H_0) = P\left(Z_0 \geq C \mid H_0 : p = p_0\right),\ \therefore C = z_\alpha .$$



At the significance level α, we reject H_0 in favor of H_a if p-value < α.

Note: (1) This large sample test is the same as the large sample test for one population mean – **but the population here is a specific one, the Bernoulli(p).

(2) We also learn another formulation for the p-value for a two-sided test. The two-sided p-value = 2 ∗ P(|Z_0| ≥ |z_0|) $= 2 * \min\{P(Z_0 \geq z_0 \mid H_a), P(Z_0 \leq z_0 \mid H_a)\}$. (This simply means that the p-value for the 2-sided test is twice the tail area bounded by z_0. You can verify the formula by examining the two cases when z_0 is positive or negative.)

Note: One can use either the original or the alternative P.Q. for constructing the large sample CI or performing the large sample test. The versions we presented here are the most common choices.

Example. A college has 500 women students and 1,000 men students. The introductory zoology course has 90 students, 50 of whom are women. It is suspected that more women tend to take zoology than men. Please test this suspicion at $\alpha = .05$.

Solution: Inference on one population proportion, large sample.

The hypotheses are:

$$H_0: p = 1/3 \text{ versus } H_a: p > 1/3$$

The test statistics is:

$$Z_0 = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} = \frac{50/90 - 1/3}{\sqrt{\frac{1/3(1-1/3)}{90}}} \approx 4.47$$

Since $Z_0 = 4.47 > Z_{\alpha=0.05} = 1.645$, we reject H_0 at the significance level of 0.05 and claim that more women tend to take zoology than men.

⑥ Exact Test:

Example. In 1000 tosses of a coin, 560 heads and 440 tails appear. Is it reasonable to assume that the coin is fair?
Let p denote the probability of obtaining head in one throw. This is sampling from the Bernoulli(p) population.
Let $X \sim \text{Binomial}(n=1000, p)$ denote the number of heads in 1000 throws – we can easily verify that X is a suffiicent statistic for p.

Solution:

H_0: $p = 1/2$ H_1: $p > 1/2$

Previously we have discussed the pivotal quantity method. Indeed, we can derive a test using only a *test statistic* $W(X_1,\ldots,X_n) = W(\mathbf{X})$, a function of the sample, that although we do not know its distribution entirely, we know its distribution exactly under the null hypothesis. This test statistic is often obtained through a point estimator (for example, X/n, the sample proportion) of the parameter of interest, or a sufficient statistic (for example, X, the total number of successes) of the parameter of interest (in the given example, it is p).

So, subsequently, we can compute the p-value, and make decisions based on the p-value.

In our example, such a test statistic could be the total number of heads when the null hypothesis is true (H_0: p = 1/2), that is:

$X_0 \sim \text{Binomial}(1000, p = 1/2)$.

The observed test statistic value is:

$$x_0 = 560$$

If the null hypothesis is true, we would expect about 500 heads. If the alternative hypothesis is true, we would expect to see more than 500 heads. The more heads we observed, the more evidence we have in supporting the alternative hypothesis. Thus by the definition, the p-value would be the probability of observing 560 or more heads, when the null hypothesis is true, i.e.

$$P(X \geq 560 | H_0: p = 1/2) = \sum_{x=560}^{1000} \binom{1000}{x} \left(\frac{1}{2}\right)^x \left(\frac{1}{2}\right)^{1000-x}$$

$$\approx 0.0000825$$

The above is what we referred to as an exact test. As you can see, in exact test, we often use the p-value method to make decisions, rather than the rejection region method.

Alternatively, using the normal approximation we have $X_0 \sim N(500, 250)$ yielding the large sample p-value of:

$$P(X_0 \geq 560 \mid H_0: p = 1/2) = P\left(\frac{X_0 - 500}{\sqrt{250}} \geq \frac{560 - 500}{\sqrt{250}} \middle| p = 1/2 \right)$$

$$\approx P(Z_0 \geq 3.795) \approx 0.0000739$$

This is the same as the large sample approximate test we had derived before using the pivotal quantity method and the central limit theorem.

Likelihood Ratio Test

(one population mean, normal population, two-sided)

1. Please derive the likelihood ratio test for H_0: $\mu = \mu_0$ versus Ha: $\mu \neq \mu_0$, when the population is normal and population variance σ^2 is known.

Solution:

For a 2-sided test of H_0: $\mu = \mu_0$ versus Ha: $\mu \neq \mu_0$, when the population is normal and population variance σ^2 is known, we have:

$\varpi = \{\mu : \mu = \mu_0\}$ and $\Omega = \{\mu : -\infty < \mu < \infty\}$

The likelihoods are:

$$L(\varpi) = L(\mu_0)$$

$$= \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x_i - \mu_0)^2}{2\sigma^2}\right) = \left(\frac{1}{2\pi\sigma^2}\right)^{\frac{n}{2}} \exp\left[-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - \mu_0)^2\right]$$

There is no free parameter in $L(\varpi)$, thus $L(\hat{\varpi}) = L(\varpi)$.

$$L(\Omega) = L(\mu)$$

$$= \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x_i - \mu)^2}{2\sigma^2}\right) = \left(\frac{1}{2\pi\sigma^2}\right)^{\frac{n}{2}} \exp\left[-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - \mu)^2\right]$$

There is only one free parameter μ in $L(\Omega)$. Now we shall find the value of μ that maximizes the log likelihood

$$\ln L(\Omega) = -\frac{n}{2}\ln\left(2\pi\sigma^2\right) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - \mu)^2 .$$

By solving $\dfrac{d \ln L(\Omega)}{d\mu} = \dfrac{1}{\sigma^2}\sum_{i=1}^{n}(x_i - \mu) = 0$, we have $\hat{\mu} = \bar{x}$

It is easy to verify that $\hat{\mu} = \bar{x}$ indeed maximizes the loglikelihood, and thus the likelihood function.

Therefore the likelihood ratio is:

$$\lambda = \frac{L(\hat{\varpi})}{L(\hat{\Omega})} = \frac{L(\mu_0)}{\max_{\mu} L(\mu)}$$

$$= \frac{L(\mu_0)}{L(\hat{\mu})} = \frac{\left(\frac{1}{2\pi\sigma^2}\right)^{\frac{n}{2}} \exp\left[-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - \mu_0)^2\right]}{\left(\frac{1}{2\pi\sigma^2}\right)^{\frac{n}{2}} \exp\left[-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - \bar{x})^2\right]}$$

$$= \exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}\left[(x_i - \mu_0)^2 - (x_i - \bar{x})^2\right]\right\}$$

$$= \exp\left[-\frac{1}{2}\left(\frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}}\right)^2\right] = \exp\left[-\frac{1}{2}(z_0)^2\right]$$

Therefore, the likelihood ratio test that will reject H_0 when $\lambda \le \lambda^*$ is equivalent to the z-test that will reject H_0 when $|Z_0| \ge c$, where c can be determined by the significance level α as $c = z_{\alpha/2}$.

You see that from the beauty contest example that to set which theory as the null hypothesis may not always follow the same rule. That is why for the normality test, we put the hypothesis that “the population from which the sample was drawn is a normal population” as the null hypothesis.

Similarly, in the beauty contest problem, it is easier to interpret when we set the null hypothesis as “Stony Brook is the most beautiful University”, versus the alternative that we are not.

2. Please derive the likelihood ratio test for H_0: $\mu = \mu_0$ versus Ha: $\mu \neq \mu_0$, when the population is normal and population variance σ^2 is unknown.

Solution: For a 2-sided test of H_0: $\mu = \mu_0$ versus Ha: $\mu \neq \mu_0$, when the population is normal and population variance σ^2 is unknown, we have:

$\varpi = \left\{\left(\mu, \sigma^2\right): \mu = \mu_0, 0 < \sigma^2 < \infty\right\}$ and

$\Omega = \left\{\left(\mu, \sigma^2\right): -\infty < \mu < \infty, 0 < \sigma^2 < \infty\right\}$

The likelihood under the null hypothesis is:

$$L(\varpi) = L\left(\mu_0, \sigma^2\right)$$

$$= \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x_i - \mu_0)^2}{2\sigma^2}\right) = \left(\frac{1}{2\pi\sigma^2}\right)^{\frac{n}{2}} \exp\left[-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - \mu_0)^2\right]$$

There is one free parameter, σ^2, in $L(\varpi)$. Now we shall find the value of σ^2 that maximizes the log likelihood

$\ln L(\varpi) = -\frac{n}{2}\ln\left(2\pi\sigma^2\right) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - \mu_0)^2$. By solving

$\frac{d \ln L(\varpi)}{d\sigma^2} = -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4}\sum_{i=1}^{n}(x_i - \mu_0)^2 = 0$, we have

$$\hat{\sigma}_{\varpi}^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \mu_0)^2$$

It is easy to verify that this solution indeed maximizes the loglikelihood, and thus the likelihood function.

The likelihood under the alternative hypothesis is:

$$L(\Omega) = L\left(\mu, \sigma^2\right)$$

$$= \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x_i - \mu)^2}{2\sigma^2}\right) = \left(\frac{1}{2\pi\sigma^2}\right)^{\frac{n}{2}} \exp\left[-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - \mu)^2\right]$$

There are two free parameter μ and σ^2 in $L(\Omega)$. Now we shall find the value of μ and σ^2 that maximizes the log likelihood

$$\ln L(\Omega) = -\frac{n}{2}\ln\left(2\pi\sigma^2\right) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - \mu)^2.$$

By solving the equation system:

$\frac{\partial \ln L(\Omega)}{\partial \mu} = \frac{1}{\sigma^2}\sum_{i=1}^{n}(x_i - \mu) = 0$ and

$$\frac{\partial \ln L(\Omega)}{\partial \sigma^2} = -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4}\sum_{i=1}^{n}(x_i - \mu)^2 = 0$$

we have $\hat{\mu} = \bar{x}$ and $\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2$

It is easy to verify that this solution indeed maximizes the loglikelihood, and thus the likelihood function.

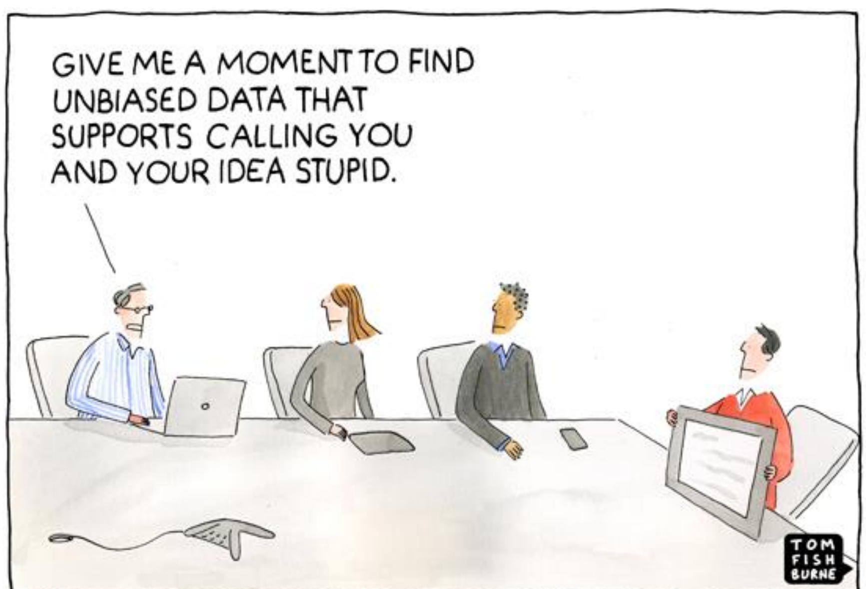
Therefore the likelihood ratio is:

$$\lambda = \frac{L(\hat{\varpi})}{L(\hat{\Omega})} = \frac{\max_{\sigma^2} L(\mu_0, \sigma^2)}{\max_{\mu,\sigma^2} L(\mu, \sigma^2)} = \frac{L(\mu_0, \hat{\sigma}^2_{\varpi})}{L(\hat{\mu}, \hat{\sigma}^2)} = \frac{\left(\frac{n}{2\pi\sum_{i=1}^{n}(x_i - \mu_0)^2}\right)^{\frac{n}{2}} \exp\left[-\frac{n}{2}\right]}{\left(\frac{n}{2\pi\sum_{i=1}^{n}(x_i - \bar{x})^2}\right)^{\frac{n}{2}} \exp\left[-\frac{n}{2}\right]}$$

$$= \left(\frac{\sum_{i=1}^{n}(x_i - \mu_0)^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2}\right)^{-\frac{n}{2}} = \left(\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2 + n(\bar{x} - \mu_0)^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2}\right)^{-\frac{n}{2}} = \left(1 + \frac{n(\bar{x} - \mu_0)^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2}\right)^{-\frac{n}{2}}$$

$$= \left(1 + \frac{(t_0)^2}{n-1}\right)^{-\frac{n}{2}}$$

Therefore, the likelihood ratio test that will reject H_0 when $\lambda \le \lambda^*$ is equivalent to the t-test that will reject H_0 when $|t_0| \ge c$, where c can be determined by the significance level α as $c = t_{n-1,\alpha/2}$.



© marketoonist.com

Likelihood Ratio Test
(one population mean, normal population, one-sided)

1. Let $f(x_i|\mu) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x_i-\mu)^2}{2\sigma^2}}, for\ i = 1, \cdots, n,$ where σ^2 is known. Derive the likelihood ratio test for the hypothesis $H_0: \mu = \mu_0\ vs. H_a: \mu > \mu_0$.

Solution:

$$\boldsymbol{\omega} = \{\mu: \mu = \mu_0\}$$
$$\boldsymbol{\Omega} = \{\mu: \mu \geq \mu_0\}$$

Note: Now that we are not just dealing with the two-sided hypothesis, it is important to know that the more general definition of $\boldsymbol{\omega}$ is the set of all unknown parameter values under H_0, while **$\boldsymbol{\Omega}$ is the set of all unknown parameter values under the union of H_0 and H_a.**

Therefore we have:

$$\begin{cases} L_\omega = f(x_1, x_2, \cdots, x_n|\mu = \mu_0) = \prod_{i=1}^{n} f(x_i|\mu = \mu_0) \\ = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x_i-\mu_0)^2}{2\sigma^2}} = (2\pi\sigma^2)^{-\frac{n}{2}} e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu_0)^2} \\ L_\Omega = f(x_1, x_2, \cdots, x_n|\mu \geq \mu_0) = \prod_{i=1}^{n} f(x_i|\mu \geq \mu_0) \\ = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x_i-\mu)^2}{2\sigma^2}} = (2\pi\sigma^2)^{-\frac{n}{2}} e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2} \end{cases}$$

Since there is no free parameter in L_ω,

$$\sup(L_\omega) = (2\pi\sigma^2)^{-\frac{n}{2}} e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu_0)^2}.$$

$$\because \begin{cases} \frac{\partial log L_\Omega}{\partial \mu} = \frac{n}{\sigma^2}(\bar{x} - \mu) \\ \frac{\partial^2 log L_\Omega}{\partial \mu^2} = -\frac{n}{\sigma^2} < 0 \end{cases},$$

$$\therefore \sup(L_\Omega) = \begin{cases} (2\pi\sigma^2)^{-\frac{n}{2}} e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\bar{x})^2}, if\ \bar{x} > \mu_0 \\ (2\pi\sigma^2)^{-\frac{n}{2}} e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu_0)^2}, if\ \bar{x} \leq \mu_0 \end{cases}.$$

$$\therefore \lambda = \frac{\sup(L_\omega)}{\sup(L_\Omega)} = \begin{cases} e^{-\frac{n}{2\sigma^2}(\bar{x}-\mu_0)^2}, if\ \bar{x} > \mu_0 \\ 1, if\ \bar{x} \le \mu_0 \end{cases}.$$

$$\therefore P(\lambda \le \lambda^* | H_0) = \alpha \Leftrightarrow P\left(Z_0 = \frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}} \ge \sqrt{-2ln\lambda^*}\middle| H_0\right) = \alpha.$$

$\therefore$ Reject H_0 in favor of H_a if $Z_0 = \frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}} \ge Z_\alpha$

2. Let $X_1, X_2, \dots X_n \sim N(\mu, \sigma^2)$, where σ^2 is known
   Find the LRT for $H_0: \mu \le \mu_0 \quad vs \quad H_a: \mu > \mu_0$

Solution:

$$\boldsymbol{\omega} = \{\mu: \mu \le \mu_0\}$$
$$\boldsymbol{\Omega} = \{\mu: -\infty < \mu < \infty\}$$

Note: Now that we are not just dealing with the two-sided hypothesis, it is important to know that the more general definition of $\boldsymbol{\omega}$ is the set of all unknown parameter values under H_0, while $\boldsymbol{\Omega}$ is the set of all unknown parameter values under the union of H_0 and H_a.

(1) The ratio of the likelihood is shown as the following,

$$\lambda = \frac{L(\widehat{\omega})}{L(\widehat{\Omega})}$$

$L(\widehat{\omega})$ is the maximum likelihood under $H_0: \mu \le \mu_0$

$$L(\widehat{\omega}) = \begin{cases} \left(\frac{1}{2\pi\sigma^2}\right)^{n/2} \exp\left(-\frac{\sum(x_i-\bar{x})^2}{2\sigma^2}\right), & \bar{x} < \mu_0 \\ \left(\frac{1}{2\pi\sigma^2}\right)^{n/2} \exp\left(-\frac{\sum(x_i-\mu_0)^2}{2\sigma^2}\right), & \bar{x} \ge \mu_0 \end{cases}$$

$L(\widehat{\Omega})$ is the maximum likelihood under $H_0\ union\ H_a$

$$L(\widehat{\Omega}) = \left(\frac{1}{2\pi\sigma^2}\right)^{n/2} \exp\left(-\frac{\sum(x_i-\bar{x})^2}{2\sigma^2}\right)$$

The ratio of the maximized likelihood is,

$$\lambda = \begin{cases} 1, \ \ \bar{x} < \mu_0 \\ exp\left(-\frac{(\sum(x_i-\mu_0)^2 - \sum(x_i-\bar{x})^2)}{2\sigma^2}\right), & \bar{x} \ge \mu_0 \end{cases}$$

$\Rightarrow$

$$\lambda = \begin{cases} 1, & \bar{x} < \mu_0 \\ e^{-\frac{n(\bar{x}-\mu_0)^2}{2\sigma^2}}, & \bar{x} \ge \mu_0 \end{cases}$$

By LRT, we reject null hypothesis if $\lambda \leq \lambda^*$. Generally, λ^* is chosen to be less than 1. The rejection region is,

$$R = \left\{X : e^{-\frac{n(\bar{x}-\mu_0)^2}{2\sigma^2}} \leq \lambda^* \right\} = \left\{X : \bar{X} \geq \mu_0 + \sqrt{-\frac{2\sigma^2}{n} \ln \lambda^*}\right\}$$

$$\therefore P(\lambda \leq \lambda^* | H_0) = \alpha \Leftrightarrow P\left(Z_0 = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \geq \sqrt{-2ln\lambda^*}\middle| H_0\right) = \alpha.$$

$\therefore$ Reject H_0 in favor of H_a if $Z_0 = \frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}} \geq Z_\alpha$

Now we have shown that the one-sided hypothesis tests for the following pairs of hypotheses are indeed, the same.

$$\begin{cases} H_0 : \mu = \mu_0 \\ H_a : \mu > \mu_0 \end{cases} \Leftrightarrow \begin{cases} H_0 : \mu \leq \mu_0 \\ H_a : \mu > \mu_0 \end{cases}$$