Recap: Basic Assumptions.

13th March' 2024

Section 6.2 – One Sample Models (6.2.1, 6.2.2)

Hypothesis tests and C.I's for μ

Set-up: Suppose $Y_1, Y_2, \ldots, Y_n \overset{iid}{\sim} N(\mu, \sigma^2)$ random sample.

Recall the joint log-likelihood:

$$\ell(\mu, \sigma^2) = -\frac{n}{2}\ln(\sigma^2) - \sum_{i=1}^{n}(y_i-\mu)^2 / 2\sigma^2$$

We want to test, $H_0: \mu = \mu_0$ ← some value.

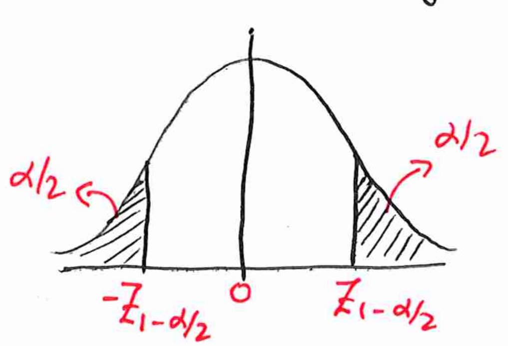
When σ^2 is known

$$\text{LRS};\ D = 2\Big[\underbrace{\ell(\hat{\mu}, \sigma^2)}_{k=1} - \underbrace{\ell(\mu_0, \sigma^2)}_{q=0}\Big]$$

...... Algebra

$$= (\hat{\mu}-\mu_0)^2 / \tfrac{\sigma^2}{n} = Z^2 \sim \chi^2_{(1)} \quad [Z \sim N(0,1)]$$

$$
\begin{aligned}
p\text{-value} &= P(D \geq d_{obs}) \\
&= P(Z^2 \geq Z^2_{obs}) \\
&= P(|Z| \geq |Z_{obs}|) \quad \leftarrow \text{2-tailed} \\
&= 2 \cdot P(Z \geq |Z_{obs}|), \text{ because } N(0,1) \text{ is symmetric.}
\end{aligned}
$$



$100(1-\alpha)\%$ C.I for μ:

$$\boxed{\hat{\mu} \pm \underbrace{Z_{1-\alpha/2}\, \sigma/\sqrt{n}}_{\text{Margin of Error}}}$$

$Z_{1-\alpha/2}$: Critical value

$\sigma/\sqrt{n}$: Standard Error

$\hat{\mu} = \bar{y}$

$\boxed{\text{When } \delta^2 \text{ unknown}}$ $H_0 : \mu = \mu_0$

$$D \equiv 2\left[\ell(\hat{\mu}, \hat{\delta}^2) - \ell(\mu_0, \tilde{\delta}^2)\right] \approx \chi^2_{(1)}$$

$k = 2$ Estimated $q = 1$

D is an 1-1 increasing function of T^2,

where, $T = \dfrac{\hat{\mu} - \mu_0}{S/\sqrt{n}} \sim t_{(n-1)}$ Student's t distribution

$S^2 = \sum_{i=1}^{n} (y_i - \hat{\mu})^2 / n-1 \rightarrow$ sample variance.

$$D = n \ln\left[1 + \frac{1}{n-1} T^2\right]$$ ★ Exercise: Derive this!

$$\begin{aligned} P\text{-value} &= P(D \geq d_{obs}) \\ &= P(T^2 \geq t^2_{obs}) \\ &= \cdots\cdots = 2P(T \geq |t_{obs}|), \end{aligned}$$ by symmetry of t-distribution.

$100(1-\alpha)\%$ C.I for μ :

$P(|T| \geq t_{1-\alpha/2, n-1}) \geq \alpha \iff \underbrace{t_{obs}(\mu_0) \leq t_{1-\alpha/2, (n-1)}}$

Solve for set of μ_0 that satisfy this.

$\Rightarrow \boxed{\hat{\mu} \pm t_{1-\alpha/2, (n-1)} \cdot S/\sqrt{n}}$ $\hat{\mu} = \bar{y}$.

Now, Let look at the Example 6.2.1 from the Complete Lecture Notes $\rightarrow$

Plotting your data before analysis important

Want to assess Normality Assumptions

n=10 is small, histograms are hard to interpret

But, nothing suspicious here.

Figure 6.1: Histogram of 10 Salaries

Q-Q plot is used for normaility check as well

If points fall on the straight line, then the distributions are the same.

Therefore, here we see that Normality assumption is fine.

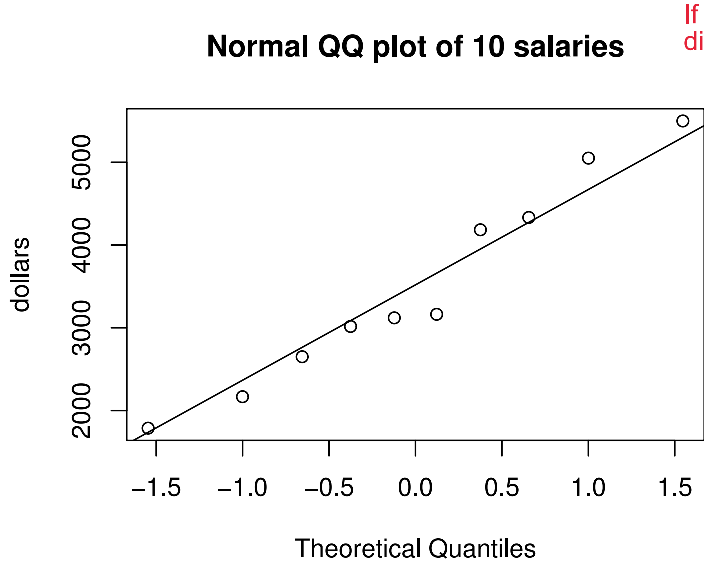


Figure 6.2: QQ plot of 10 Salaries

$\hookrightarrow$ Since, 95% C.I $\Rightarrow$ $\alpha = 0.05$

Because, δ^2 is not assumed known, we will use the $t_{(n-1)}$ version of the confidence interval formula.

$$\hat{\mu} = \bar{y} = \$3496.90$$

$$s = \sqrt{\frac{\sum_{i=1}^{n}(y_i - \hat{\mu})^2}{n-1}} = \$1224.116$$

percentile of interest

$$t_{1-\alpha/2,(n-1)} = t_{0.975(9)} = 2.262$$ } in R, qt(0.975, 9)

Degrees of freedom.

$$\hookrightarrow \hat{\mu} \pm t_{0.975(9)} \, s/\sqrt{n} = [\$2621.22, \$4372.58]$$

Note:- Consider $X \sim N(0,1)$, and the following equation:

$$P = P(X \le X_{obs})$$

in R, this is saying:

[P = pnorm(Xobs, mu=0, sd=1)]

and also,

[Xobs = qnorm(p, mu=0, sd=1)]

So, [q() functions return quantiles, given probabilities.]

and, [p() functions return probabilities, given quantiles.]

i.e., $\underset{X_{obs}}{}$ pnorm(0, 0, 1) = 0.5

qnorm(0.5, 0, 1) = 0

p

Let's look at the Example 6.2.2 from the Complete Lecture Notes ...

↪ The likelihood ratio test for testing

$H_0 : \mu = \$3000$ is

$$D = n \ln\left[1 + \frac{1}{n-1} T^2\right], \quad T = \underbrace{\frac{\bar{y} - 3000}{s/\sqrt{10}}}$$

plug in $\bar{y}$ and s from Ex: 6.2.1 to get t_{obs}.

$$\begin{aligned} p\text{-value} &= P[D \geq d_{obs}] \\ &= 2P(T \geq |t_{obs}|) \\ &= 2P(T \geq 1.2836) \\ &> 2(0.1), \text{ where } 0.1 \text{ was found using t-tables} \\ &= 0.2. \end{aligned}$$

But, before jumping into calculation,

Ask yourself, if your confidence interval from Ex: 6.2.1 tells us anything about the p-value of this test,

$$95\% = [\$2621.22, \$4372.58]$$

So, ... ?