

Stat 8003, Homework 7

Group G: `sample(c("David" , "Andrew", "Salam"))`

October 16, 2014

Question 7.1. We want to know the mean percentage of butterfat in milk produced by a farm by sampling multiple loads of milk. Previous records indicate the average percent butterfat in milk is 3.35 and the standard deviation among loads is 0.15. Now we hope to detect a change of the percent butterfat in milk.

- (a) Find the rejection region at the significance level $\alpha = 0.05$;
- (b) Suppose 100 loads of milk are sampled. What is the power for the test for detecting a change of the mean to 3.40.
- (c) Plot the power as a function of the absolute value of the change in the mean over the standard deviation (which is $|\mu_1 - \mu_0|/\sigma$).
- (d) Now we hope to detect a change of the percent butterfat in milk to 3.40 with a power 0.8. How many loads do we need to sample?

Answer:

- (a) Let $r.v.$ X be the percentage of butterfat in milk. We have

$$X \sim N(\mu = 3.35, \sigma = 0.15)$$

Or

$$Z = \frac{X - 3.35}{0.15} \sim N(0, 1)$$

The null hypothesis is that there is no change in butterfat; the alternative hypothesis is that there is:

$$H_0 : \mu = 3.35$$

$$H_a : \mu \neq 3.35$$

The rejection region of this two sided test is given by

$$\begin{aligned}
\mathcal{R} &= \left\{ |Z| > z_{\alpha/2} \right\} \\
&= \left\{ Z < z_{0.975} \right\} \cup \left\{ Z > z_{0.025} \right\} \\
&= \left\{ \frac{X - 3.35}{0.15} < -z_{0.025} \right\} \cup \left\{ \frac{X - 3.35}{0.15} > z_{0.025} \right\} \\
&= \left\{ X < 3.35 - 0.15 \times 1.96 \right\} \cup \left\{ X > 3.35 + 0.15 \times 1.96 \right\} \\
&= \left\{ X < 3.056 \right\} \cup \left\{ X > 3.644 \right\}
\end{aligned}$$

- (b) We want the probability of rejecting the null when H_a is true (specifically, when $\mu = 3.40$). With 100 loads of milk sampled, $\bar{X} \sim N(\mu, \sigma/\sqrt{100})$.

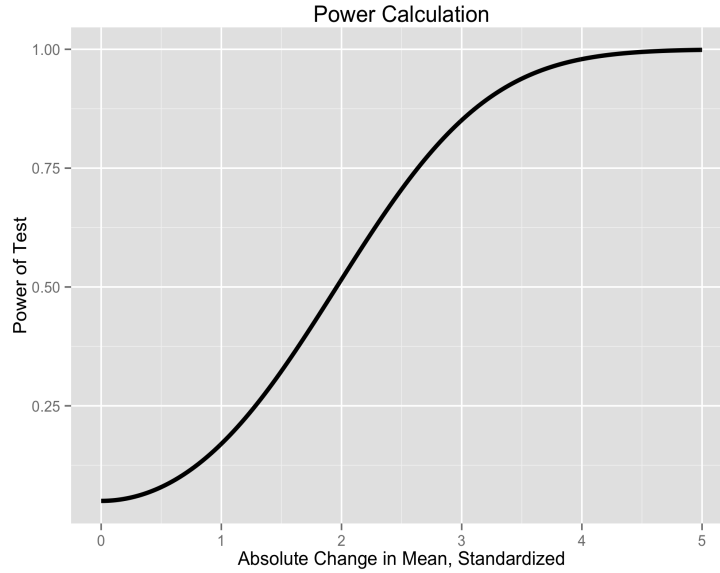
The probability of rejecting H_0 when in fact $\mu = 3.40$:

$$\begin{aligned}
1 - \beta &= P(\mathcal{R} \mid \mu = 3.4) = P\left(\left|\frac{\bar{X} - 3.35}{0.015}\right| > z_{0.025}\right) \\
&= P\left(\left|\frac{\bar{X} - 3.4 + (3.4 - 3.35)}{0.015}\right| > 1.96\right) \\
&= P\left(\left|Z + \frac{(3.4 - 3.35)}{0.015}\right| > 1.96\right) \\
&= P\left(Z > 1.96 - \frac{(3.4 - 3.35)}{0.015}\right) + P\left(Z < -1.96 - \frac{(3.4 - 3.35)}{0.015}\right) \\
&= P(Z > -1.373333) + (\text{negligible quantity}) \\
&= 1 - \text{pnorm}(-1.373333)
\end{aligned}$$

$\text{power} = 1 - \beta = 0.915$

- (c) We now we express the power in terms of the absolute change in the mean over its standard deviation ($|\mu_1 - \mu_0|/\sigma$) and plot the result.

$$\begin{aligned}
1 - \beta &= P\left(\left|\frac{\bar{X} - \mu_1 + (\mu_1 - \mu_0)}{\sigma}\right| > z_{0.025}\right) \\
&= P\left(\left|Z + \frac{\mu_1 - \mu_0}{\sigma}\right| > 1.96\right) \\
&= P\left(Z > 1.96 - \frac{|\mu_1 - \mu_0|}{\sigma}\right) + P\left(Z < -1.96 - \frac{|\mu_1 - \mu_0|}{\sigma}\right) \\
&= 1 - \text{pnorm}(1.96 - \text{xaxis}) + \text{pnorm}(-1.96 - \text{xaxis})
\end{aligned}$$



(d) Using the result obtained in class with $1 - \beta = 0.8$, we get:

$$n = \frac{\sigma^2}{\delta^2} (z_{\frac{\alpha}{2}} + z_{\beta})^2$$

$$= \frac{0.15^2}{(3.40 - 3.35)^2} (z_{0.025} + z_{0.2})^2$$

$$n = 71$$

Question 7.2. The relative rotation angle between the L2 and L3 lumbar vertebrae is defined as the acute angle between posterior tangents drawn to each vertebra on a spinal X-ray. When this angle is too large the patient experiences discomfort or pain. Chiropractic treatment of this condition involves decreasing this angle by applying (nonsurgical) manipulation or pressure. Harrison et al. (2002) propose one such particular treatment. They measured the angle on both pre- and post-treatment X-rays from a random sample of 48 patients. At $\alpha = 0.05$, test whether the mean post-treatment angle is less than the mean angle prior to treatment.

Answer:

We are comparing the means of pre- and post-treatment angles. Let the null hypothesis be that there is no change in the mean angle between pre- and post-treatments. We are testing for the left-sided alternative that there's a decrease in the angle.

$$H_0 : \mu_{\text{post}} = \mu_{\text{pre}}$$

$$H_a : \mu_{\text{post}} < \mu_{\text{pre}}$$

We first check whether the variances of the two groups are equal ($\sigma_{\text{post}} = \sigma_{\text{pre}}$) using the F -test.

The F -statistic with 47 and 47 degrees of freedom returned by the F -test is 1.3369 and the p -value is 0.1615073. This p -value is too high offering us little evidence that the variances are different. We thus *accept* the hypothesis that the variances of the two groups are equal.

We use R's paired T-test to see whether the mean post-treatment angle is less than the pre-treatment mean angle:

```
t.test( har$Post, har$Pre, var.equal=TRUE,
        alternative="less", paired=TRUE)
```

This outputs $p\text{-value} = 0.0004353$ and we therefore *reject* the null hypothesis that the means are equal.

Accept H_1 : The post-treatment mean is smaller

Question 7.3. This problem will guide you in demonstrating the multiplicity issues in multiple hypothesis testing using simulation. Let X_{ij} be the data modeled as

$$X_{ij} \stackrel{iid}{\sim} N(\theta_j, 1), \quad i = 1, 2, \dots, n \quad j = 1, 2, \dots, p. \quad (1)$$

Consider testing p hypotheses $H_{0j} : \theta_j = 0$ vs $H_{1j} : \theta_j \neq 0$ where $j = 1, 2, \dots, p$. Define the family-wise error rate (FWER) as

$$FWER = P(\text{at least one (including one) false rejection}).$$

Set the sample size $n = 50$ and $\alpha = 0.05$.

- (a) Set $p = 1$, $\theta_j = 0$, $\forall j$. Generate X_{ij} according to (1) and use p -value to test the hypothesis at α -level. Replicate this step 1000 times to get the simulated FWER.
- (b) Set $p = 10$, $\theta_j = 0$, $\forall j$. Generate X_{ij} according to (1) and use p -value to test 10 hypothesis at α -level. Replicate this step 1000 times to get the simulated FWER.
- (c) Set $p = 100$, $\theta_j = 0$, $\forall j$. Generate X_{ij} according to (1) and use p -value to test 100 hypothesis at α -level. Replicate this step 1000 times to get the simulated FWER.
- (d) Set $p = 100$, $\theta_j = 0$, $\forall j$. Generate X_{ij} according to (1) and use p -value to test 100 hypothesis using Bonferroni's correction by setting the significance level at α/p for each hypothesis. Replicate this step 1000 times to get the simulated FWER.

Answer:

To answer this question, we wrote an R function that we use for parts (a) through (d). This function tests any number of hypotheses on a batch of 50 random numbers drawn from the standard normal distribution, where each hypothesis being tested is whether the mean of these 50 numbers equals zero, at a specified level of significance α (please see attached code for details).

Letting $X_1, X_2, \dots, X_{n=50}$ be a random sample from a $N(\theta, 1)$ population, we reject $H_0 : \theta = 0$ in favor of $H_1 : \theta \neq 0$ whenever

$$|Z| = \left| \frac{\bar{X} - \theta}{1/\sqrt{n}} \right| = \sqrt{50} |\bar{X}| \geq z_{0.025} = 1.96,$$

i.e. whenever

$$\{\bar{X} \leq -0.27719\} \cup \{\bar{X} \geq 0.27719\}.$$

- (a) We first test a single hypothesis $H_0 : \theta = 0$ on one batch of 50 standard normals and *accept* it.

We then repeat this step 1000 times and get the result that about 0.044 of these hypotheses are rejected. (Of course this number varies with each run of the simulation, but it remains close to 0.05, as expected.)

- (b) We now test 10 hypotheses for each of the 1000 simulations. For each simulation run, we determine whether any of the 10 hypotheses have been rejected. Whenever at least one of the hypotheses is rejected, it contributes to the family wise error rate (FWER).

At the end of this process, we find that FWER = 0.39 - roughly a ten-fold increase from the result in part (a). (Again, this number varies slightly from one run to another).

- (c) We rerun part (b), this time with 100 hypotheses and find that the FWER shoots up to 0.995.
- (d) Finally, using the suggested Bonferroni's correction, we set $\alpha = 0.05/100$ and rerun part (c). We see that we are now down to FWER = 0.033.