Homework 6 Solution

1 Frequentist Methods

1.1 Z test

H_0: $\mu = 0$

H_A: $\mu > 0$

The test statistic is

$$Z = \frac{\bar{x} - 0}{s/\sqrt{n}} = \frac{33.7}{66.2/\sqrt{11}} = 1.69,$$

which we assume to be distributed as a standard normal random variable.

An extreme value is large values of Z, because the alternate hypothesis is greater than; it's a one-sided test.

The p-value is then

$$\Pr(Z > 1.69) \approx 0.0455.$$

1.2 T test

This test is the same as above, except we now assume that the test statistic is distributed with a T distribution with $n - 1 = 10$ degrees of freedom. The p-value is then

$$\Pr(T_{10} > 1.69) \approx 0.0609.$$

1.3 Sign test

H_0: $m = 0$, where m is the median

H_A: $m > 0$

The test statistic Y is the number of samples with differences greater than the median for the null hypothesis (that is, 0); there are 7. This has a $Y \sim Bin(11, 0.5)$ distribution.

An extreme value is to have a large number of differences greater than the null hypothesis median, because the alternate hypothesis is greater than; it's a one-sided test.

So the p-value is

$$\Pr(Y \geq 7) = 0.2744 \quad \text{from the table of page 651.}$$

1.4 Sign rank test

For this test, the hypotheses are as in the sign test; to find the test statistic, we must rank the data points by their absolute value:

Rank	1	2	3	4	5	6	7	8	9	10	11
Data	−20	24	−33	−36	38	62	−70	72	101	106	127

The test statistic is the sum of ranks of the negative numbers (1+3+4+7=15); this has a distribution as on page 664.

An extreme test statistic is to have few negative numbers, as that means that the median is positive, so looking this up in the table, the p-value is 0.062.

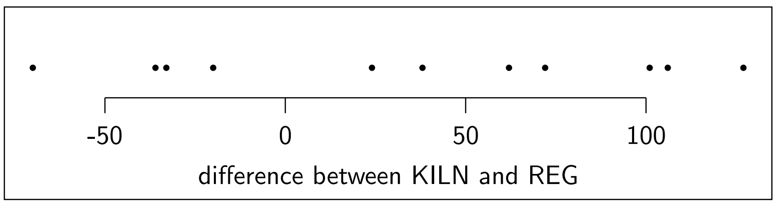
1.5 Different p-values

We have different p-values because each of these tests has different assumptions, which changes the distribution of the test statistic. The p-value is the probability of getting extreme values, which depends on what the distribution of test statistic is assumed to be.

1.6 So which tests are appropriate?

The Z test is not appropriate; for it we need a large enough sample size for the normal approximation given by the CLT to hold. A rule of thumb for this is 30; we have only 11.

The T test is appropriate for small sample sizes if the data is distributed close to normally. Below is a plot of the 11 data points; they don't have any obvious skew so I'd say they're probably close to normal. You may disagree, or you may have used a Q-Q plot that you've learned about in another class. If you agree that they're close to normal, this test is appropriate, otherwise it isn't.



The sign test has no assumptions about the data, so it is appropriate to use.

The sign rank test only requires that the data be symmetrically distributed around the median; from the plot above, we see that it is approximately symmetrical, so this test is also appropriate to use.

2 Bayesian methods

We now find some probabilities using Bayesian methods; these are not p-values, as they are not probabilities about extreme data under the null. Instead they are probabilities about the null hypothesis itself (is the difference really positive?). Frequentist methods can't answer this question; that's why frequentist methods use p-values. Bayesian methods can, but they pay the price of having to force a prior onto the parameter of interest.

2.1 Finding the posterior

The given prior is $\mu \sim N(0, 20^2)$, and the data we assume is distributed $N(\mu, 66.2^2)$. So we can use the calculation from p. 358 to find the posterior; it's normally distributed with mean and variance as given below, where ν and tau are the prior mean and standard deviation, 0 and 20; σ is the population standard deviation, 66.2; n is the sample size, 11; and $\bar{x}$ is the sample mean, 33.7.

$$E(\mu|\boldsymbol{x}) = \frac{1/\tau^2}{1/\tau^2 + n/\sigma^2}\nu + \frac{n/\sigma^2}{1/\tau^2 + n/\sigma^2}\bar{x} = \frac{11/66.2^2}{1/20^2 + 11/66.2^2}33.7 \approx 16.9$$

$$Var(\mu|\boldsymbol{x}) = \frac{\sigma^2\tau^2}{\sigma^2 + n\tau^2} = \frac{66.2^2 20^2}{66.2^2 + 11(20^2)} \approx 14.13^2$$

so the posterior distribution is $\mu|\boldsymbol{x} \sim N(16.9, 14.13^2)$.

2.2 Probability of null hypothesis

$$\Pr(\mu > 0|\boldsymbol{x}) = \Pr\left(\frac{\mu - 16.9}{14.13} > \frac{-16.9}{14.13}\right) = \Pr(Z > -1.196) \approx 0.88$$

2.3 Probability of a practically significant difference

$$\begin{aligned}\Pr(-10 < \mu < 10|\boldsymbol{x}) &= \Pr\left(\frac{-10 - 16.9}{14.13} < Z < \frac{10 - 16.9}{14.13}\right) \\ &= \Pr(-1.9 < Z < -0.49) \approx 0.28\end{aligned}$$