

Statistics 3858b Assignment 4

Handout March 22, 2017 ; Due date: TBA

These problems use some data from the text.

1. In the two sample normal case (see handout from class) consider the hypothesis test of

$$H_0 : \mu_X - \mu_Y = \delta_0 \text{ versus } H_A : \mu_X - \mu_Y > \delta_0$$

where δ_0 is a specific number.

- (a) Derive the GLR (generalized likelihood ratio) test.
- (b) Show the rejection region is of the form

$$R = \left\{ \mathbf{x}, \mathbf{y} : \frac{(\bar{x} - \bar{y} - \delta_0)}{\sqrt{S_n^2 \left(\frac{1}{n} + \frac{1}{m}\right)}} > c \right\}$$

for some appropriate constant c .

2. Use the data from Problem 11.40 g. The field present data is a sample X_1, \dots, X_n from one population distribution, say F , and field absent is a sample Y_1, \dots, Y_m from a different population distribution, say G .

- (a) Suppose the F and G are normal distributions, with means μ_1, μ_2 and equal variance. Recall from class we derived the GLR test of $H_0 : \mu_1 - \mu_2 = \delta$ versus the alternative $H_A : \mu_1 - \mu_2 \neq \delta$ and showed it is equivalent to the so called Student's t statistic with pooled variance.

Base the appropriate tests and confidence intervals on the studentized random variable

$$T = \frac{(\bar{X} - \bar{Y} - \delta)}{\sqrt{S_p^2 \sqrt{\frac{1}{n} + \frac{1}{m}}}}$$

where \bar{X} is the sample mean r.v. and S_p^2 is the pooled sample variance r.v.

Carry out the test of $H_0 : \mu_1 = \mu_2$ versus the alternative $H_A : \mu_1 \neq \mu_2$ at level $\alpha = .05$. Also give a 95% confidence interval for $\mu_1 - \mu_2$.

- (b) Using this same test T above, use the nonparametric bootstrap method to give the 95% confidence interval for $\delta = E_F(S) - E_G(Y) = \mu_1 - \mu_2$. Do this using 2999 bootstrap replicates. Give the .025 and .975 quantiles, as well as the confidence interval.
3. Use the data for Ozone group in 11.6, Question 35. It is a sample of size 22. Consider methods to obtain the confidence interval for $\mu = \mu(f)$, the population mean where F is the cdf and f pdf of the population distribution.

$$\mu(f) = \int_{-\infty}^{\infty} xf(x)dx .$$

This notation is intended to show that the population mean depends of the Base the confidence interval on

$$W = \frac{\sqrt{n}(\bar{X} - \mu(f))}{\sqrt{S^2}} . \quad (1)$$

A test of the variance parameter, $H_0 : \sigma^2 = \sigma_0^2$ versus $\sigma^2 \neq \sigma_0^2$ can be based on a statistic

$$V = \frac{(n-1)S^2}{\sigma_0^2} .$$

Use this to obtain a confidence interval for σ^2 .

Below all confidence intervals will be 95% confidence intervals so the quantiles needed are the .025 and .975 quantiles. For the bootstrap methods use 2999 bootstrap replicates.

- Use the R package boot to obtain the .025 and .975 quantiles of (1).
- Use the student's t distribution to obtain the .025 and .975 quantiles of (1).
- Fit a parametric normal model to the data. Use the parametric bootstrap to obtain the .025 and .975 quantiles for (1).
- Give the 95% confidence intervals for μ using the three methods above.
- Use the nonparametric bootstrap method to obtain the .025 and .975 quantiles of V .
- Use the normal model and the bootstrap method to give the corresponding confidence intervals for σ^2 .

Notice The final exam is on April 10, 9 AM. See the course web page for the room.