

# STA 260s20 Assignment Seven: More hypothesis testing, still Part One<sup>1</sup>

The following homework problems are not to be handed in. They are preparation for Quiz 7 (Week of March 16th) and Term Test 2. **Please try each question before looking at the solution.** Use the formula sheet.

1. Based on a random sample from a Bernoulli distribution, testing a null hypothesis about  $\theta$  yields  $Z = -1.82$ .
  - (a) For  $H_0 : \theta = \theta_0$  versus  $H_1 : \theta \neq \theta_0$ 
    - i. What is the  $p$ -value?
    - ii. Do you reject  $H_0$  at  $\alpha = 0.05$ ?
    - iii. What do you conclude? Pick one:  $\theta < \theta_0$      $\theta > \theta_0$      $\theta = \theta_0$ .
  - (b) For  $H_0 : \theta \leq \theta_0$  versus  $H_1 : \theta > \theta_0$ 
    - i. What is the  $p$ -value?
    - ii. Do you reject  $H_0$  at  $\alpha = 0.05$ ?
    - iii. What do you conclude? Pick one:  $\theta \leq \theta_0$      $\theta > \theta_0$  .
  - (c) For  $H_0 : \theta \geq \theta_0$  versus  $H_1 : \theta < \theta_0$ 
    - i. What is the  $p$ -value?
    - ii. Do you reject  $H_0$  at  $\alpha = 0.05$ ?
    - iii. What do you conclude? Pick one:  $\theta \geq \theta_0$      $\theta < \theta_0$ .
  
2. For  $i = 1, \dots, n$ , let  $Y_i = \beta x_i + E_i$ , where
 

$x_1, \dots, x_n$  are fixed, known constants

$E_1, \dots, E_n$  are independent and identically distributed  $\text{Normal}(0, \sigma^2)$  random variables. The parameter  $\beta$  is unknown, but we will pretend that somehow the variance  $\sigma^2$  is *known*.

  - (a) What is the distribution of  $Y_i$ ? Include the parameters. Are the  $Y_i$  independent?
  - (b) Derive a formula the MLE of  $\beta$ . Don't bother with the second derivative test.
  - (c) What is the distribution of  $\hat{\beta}$ ? Include the parameters. Show some work.
  - (d) Standardizing  $\hat{\beta}$ , obtain a  $Z$  statistic for testing  $H_0 : \beta = \beta_0$  against  $H_1 : \beta \neq \beta_0$ .
  - (e) Suppose  $\sigma^2 = 4$  and the null hypothesis is  $\beta = 0$ . Calculate the test statistic for these data:
 

$x$	1.00	1.00	2.00	2.00	3.0	3.00
$y$	4.29	-0.59	4.66	1.71	2.5	2.98
  - (f) What is the  $p$ -value? Show some work, including a picture.
  - (g) Do you reject  $H_0$  at  $\alpha = 0.01$ ? Answer Yes or No.
  - (h) What do you conclude? Pick one:  $\beta = 0$      $\beta > 0$      $\beta < 0$ .

---

<sup>1</sup>Copyright information is at the end of the last page.

3. Here is an example from an earlier assignment. In a taste test, customers are choosing between two coffees. We adopt a Bernoulli model for the probability of choosing Coffee A, and we plan to test  $H_0 : \theta = \frac{1}{2}$  against  $H_1 : \theta \neq \frac{1}{2}$  at the 0.05 significance level using the test statistic

$$Z_n = \frac{\sqrt{n}(\bar{X}_n - \theta_0)}{\sqrt{\theta_0(1 - \theta_0)}}.$$

- (a) Derive a general formula for the power of the test of  $H_0 : \theta = \theta_0$  when the true parameter value is  $\theta$ . Use  $\Phi(\cdot)$  to denote the cumulative distribution function of a standard normal.
- (b) Verify that your formula yields  $\alpha$  when  $\theta = \theta_0$ .
- (c) Suppose that in the taste test,  $\theta_0 = 0.5$ , and the true value of  $\theta = 0.45$ , so that in the population of consumers being sampled, 45% would prefer Coffee A. What is the approximate power of the test for  $n = 100$  participants?
- (d) What is the approximate power for  $n = 783$ ? One of the numbers is off the charts; do your best.
4. Before actually carrying out a study, it's a good idea to decide the sample size on some rational basis. For all statistical tests used in practice, the power gets closer and closer to one as the sample size increases, as long as the null hypothesis is false. That is, the test becomes more and more sensitive, and more able to detect the truth. This is the key to using statistical power to decide on a sample size. Choose a particular way the null hypothesis might be wrong, and also choose a desired power – that is, a desired probability of correctly rejecting the null hypothesis for that particular scenario. Then increase the sample size until the desired power is attained.

Accordingly, please return to the appetizing example of the rat hairs in peanut butter. We are confident that the number of rat hairs in a jar of peanut butter follows a Poisson distribution, and there is a government standard that says the expected number of rat hairs in a 500g jar may be no more than 8. So using the traditional  $\alpha = 0.05$  significance level, we will test  $H_0 : \lambda \leq 8$  against  $H_1 : \lambda > 8$  with the test statistic

$$Z_n = \frac{\sqrt{n}(\bar{X}_n - \lambda_0)}{\sqrt{\lambda_0}}.$$

Suppose the true expected number of rat hairs is nine. What sample size is required for the probability of rejecting  $H_0$  to be at least 0.80? The answer is an integer. Show your work. Hint: You can directly solve for  $n$ , and then increase your answer to the next integer.

5. Let  $X_1, \dots, X_n$  be a random sample from a  $\text{Normal}(\mu, 1)$  distribution. We plan to test  $H_0 : \mu \geq 0$  against  $H_1 : \mu < 0$  with a  $Z$  test, using the usual  $\alpha = 0.05$  significance level. We want to be able to reject  $H_0$  with probability 0.80 when the true value of  $\mu$  equals minus one. What sample size is required. That is, what's the smallest sample size that will get the job done?

6. Let  $X_1, \dots, X_n$  be a random sample from a  $\text{Normal}(\mu, \sigma^2)$  distribution, and suppose we are testing  $H_0 : \sigma^2 = \sigma_0^2$  with a chi-squared test. Prove that the null hypothesis is rejected at significance level  $\alpha$  with a two-sided test if and only if the  $(1 - \alpha)100\%$  confidence interval does not contain  $\sigma_0^2$ . This example shows that as one student pointed out after lecture, the correspondence between tests and confidence intervals does not depend on the distribution of the test statistic being symmetric.
7. Suppose that a null hypothesis is to be rejected when some test statistic  $T$  exceeds a critical value  $c$ . Also suppose that  $T$  is a continuous random variable with cumulative distribution function  $F(t)$  when the null hypothesis is true. The distribution of  $T$  is supported on a single interval (which might be infinite), guaranteeing that  $F(t)$  has a unique inverse.
- The  $p$ -value is a random variable, a function of  $T$ . Write a formula for the  $p$ -value.
  - Show that when the null hypothesis is true, the  $p$ -value has a uniform distribution on the interval from zero to one.
8. The objective of *meta-analysis* is to pool the results of several studies on the same or related topics, and come to a single conclusion. For example, there might be multiple clinical trials of a new drug, with some results positive, and some inconclusive or negative. Usually the original data are not available, just the published statistical results.

Suppose we have a collection of  $m$  studies carried out by different investigators, each of which did some statistical test and came up with a  $p$ -value. Our null hypothesis is that the null hypothesis was true in *all* of the studies.

- What is the joint distribution of  $P_1, \dots, P_m$  under our null hypothesis that no effect was present in any of the studies?
- What is the distribution of  $Y_i = -2 \ln(P_i)$ ? Show your work.
- What is the distribution of  $W = \sum_{i=1}^m Y_i$ ? No proof is required. Just write down the answer.
- Why would a large value of  $W$  be surprising if all  $m$  null hypotheses were true?
- What is the critical value of  $W$ ? It's a quantile of the chi-squared distribution.
- Suppose that 6 studies on a particular topic yield the following  $p$ -values: 0.255, 0.065, 0.268, 0.044, 0.080, 0.135. Notice that only one test rejected the null hypothesis at  $\alpha = 0.05$ , but all the  $p$ -values are on the small side. It's really hard to judge whether these results provide support for the idea being tested. Do the meta-analysis test. What do you conclude?

9. Here is the standard model for testing equality of  $k$  means. There are  $k$  “groups” of observations (data values), maybe from  $k$  different experimental treatments. There are  $n_j$  observations in group  $j$ , for  $j = 1, \dots, k$ . The total number of observations is  $n = \sum_{j=1}^k n_j$ .

Independently for  $j = 1, \dots, k$  and  $i = 1, \dots, n_j$ ,  $X_{i,j} \sim N(\mu_j, \sigma^2)$ . We seek to test  $H_0 : \mu_1 = \mu_2 = \dots = \mu_k$  at significance level  $\alpha$ . We will use  $F = \frac{W_1/\nu_1}{W_2/\nu_2}$ .

- (a) Let  $W_2 = \sum_{j=1}^k \frac{(n_j-1)S_j^2}{\sigma^2}$ .
- Just to establish that you understand the notation, give a formula for  $S_j^2$ .
  - What is the distribution of  $W_2$ , and why?
  - Verify that  $W_2 = \frac{1}{\sigma^2} \sum_{j=1}^k \sum_{i=1}^{n_j} (X_{i,j} - \bar{X}_j)^2$ .
- (b) Show  $\sum_{j=1}^k \sum_{i=1}^{n_j} (X_{i,j} - \bar{X})^2 = \sum_{j=1}^k n_j (\bar{X}_j - \bar{X})^2 + \sum_{j=1}^k \sum_{i=1}^{n_j} (X_{i,j} - \bar{X}_j)^2$ , where  $\bar{X} = \sum_{j=1}^k \left(\frac{n_j}{n}\right) \bar{X}_j$ .
- (c) Verify that  $\bar{X}$  is the sample mean of all the observations.
- (d) If  $H_0 : \mu_1 = \mu_2 = \dots = \mu_k$  is true, say if all the  $\mu_j = \mu$ , what is the distribution of  $\frac{1}{\sigma^2} \sum_{j=1}^k \sum_{i=1}^{n_j} (X_{i,j} - \bar{X})^2$ ? Why?
- (e) Let  $W_1 = \frac{1}{\sigma^2} \sum_{j=1}^k n_j (\bar{X}_j - \bar{X})^2$ . What is the distribution of  $W_1$  if  $H_0$  is true? Why?
- (f) How do you know  $W_1$  and  $W_2$  are independent?.
- (g) Write a formula for the  $F$  statistic. What are the degrees of freedom?
- (h) Here’s some vocabulary and notation.
- The *Total Sum of Squares*, denoted by  $SSTO = \sum_{j=1}^k \sum_{i=1}^{n_j} (X_{i,j} - \bar{X})^2$  is the sum of squared deviations of all the observations from the overall mean.
  - The *Sum of Squares Within Groups* is denoted by  $SSW = \sum_{j=1}^k \sum_{i=1}^{n_j} (X_{i,j} - \bar{X}_j)^2$ . It’s the amount of variation still unexplained once you know group membership.
  - The *Sum of Squares Between Groups* is denoted by  $SSB = \sum_{j=1}^k n_j (\bar{X}_j - \bar{X})^2$ . It’s the (weighted) sum of squared deviation of the group means from the overall mean. Since  $SSTO$  is total variation and  $SSW$  is variation that is still unexplained after you know group membership,  $SSB$  must be the amount of variation that is explained by group membership.
  - $R^2 = \frac{SSB}{SSTO}$  is the *proportion* of variation in the observations that is explained by group membership. It’s an index of how strongly the  $X_{i,j}$  values are related to group membership.

Finally, here is a nice easy question about all this. Show  $F = \left(\frac{n-k}{k-1}\right) \left(\frac{R^2}{1-R^2}\right)$ .

- (i) Frequently, reports will give a table of  $n$ 's, means and standard deviations, but no hypothesis test. Fortunately the information in the table is all you need. You may have wondered which U of T campus hands out the highest marks. Here are some sample data collected long ago. Normality is not a bad assumption.

Campus	$n$	Mean	Standard Deviation
SG	117	2.81	0.589
UTM	32	2.70	0.478
UTSC	51	2.61	0.553

- i. Carry out the  $F$  test using the  $\alpha = 0.05$  significance level. The critical value is 3.04. I got it using R, not the awful tables of the  $F$  distribution. What do you conclude?
- ii. What is  $R^2$ ? Give a number. Does this seem like a strong relationship between campus and GPA?

---

This assignment was prepared by [Jerry Brunner](#), Department of Mathematical and Computational Sciences, University of Toronto. It is licensed under a [Creative Commons Attribution - ShareAlike 3.0 Unported License](#). Use any part of it as you like and share the result freely. The L<sup>A</sup>T<sub>E</sub>X source code is available from the course website:

<http://www.utstat.toronto.edu/~brunner/oldclass/260s20>