Quiz 4 for STA 250/MTH 342 – Fall 2017

Time: 25 mins. Closed book. Closed notes. Please show your work to get credits!

1. True or False:

T (a) The p-value is a random variable.

F (b) If the p-value is .03, the corresponding test will reject at level .01.

F (c) The p-value of a test is the probability that the null hypothesis is true.

F (d) In testing a simple hypothesis versus a simple hypothesis via the likelihood ratio, the p-value equals the likelihood ratio.

F (e) If the null hypothesis is true, we expect the p-value to be large. That is, it will take values > 0.5 more often than those ≤ 0.5.

2. Suppose n students from School District 1 and also n students from School District 2 are randomly sampled, and we record their SAT scores. Let $X_1, X_2, \ldots, X_n$ be the scores of students from District 1 and $Y_1, Y_2, \ldots, Y_n$ the scores from District 2. Suppose that for each district, the scores are distributed as independent normals, with mean μ_1 and μ_2 respectively but the same variance σ^2. If $n = 20$, $\bar{X} = 1500$, $\bar{Y} = 1510$, $s_X^2 = 100$, and $s_Y^2 = 95$, where s_X^2 and s_Y^2 are the sample variances. We are interested in testing

$$H_0 : \mu_1 = \mu_2 \quad \text{vs} \quad H_1 : \mu_1 \neq \mu_2.$$

(a) What is the pooled sample variance s^2_{pooled}? $S_{pooled}^2 = \frac{(n-1)S_X^2 + (n-1)S_Y^2}{2(n-1)} = \frac{1}{2}(S_X^2 + S_Y^2) = 97.5$

(b) What is the appropriate t-test statistic? Calculate its value for this problem. What is its exact sampling distribution under H_0? Independent samples $\Rightarrow$ 2-sample t-test

$T = \frac{\bar{X} - \bar{Y}}{S_{pooled} \cdot \sqrt{\frac{1}{n} + \frac{1}{n}}} = \frac{1500 - 1510}{\sqrt{97.5 \cdot \frac{2}{20}}} \approx -3.2$

(c) Do you expect this sampling distribution to be similar to the standard Gaussian distribution? Why or why not? Yes. $t_{38} \approx N(0,1)$ b/c $t_{38} \sim \frac{N(0,1)}{\sqrt{\chi^2_{38}/38}}$

(d) Now if you accidentally treated the two samples as paired data. That is if you treat (X_i, Y_i) as a correlated pair, what is the t-test statistic you would use to test H_0? What would be the corresponding rejection region for the test at level α.

(e) (Extra credit) Because the data are in fact unpaired, is the test in part (d) still valid in the sense that its Type I error is indeed still α?

(d) Paired 2-sample t-test $U_i = X_i - Y_i$ $S_u^2 = \Sigma(U_i - \bar{U})^2/(n-1)$

$t_{paired} = \bar{U}/S_u \cdot \sqrt{\frac{1}{n}}$ Reject when $|t_{paired}| > F_{t_{19}}^{-1}(1 - \frac{\alpha}{2})$

(e) It's still valid, as t_{paired} will still have t_{19} distn under the null.