Solution To the Final Exam Statistics 101 Spring 2004

May 7, 2004

1. A. $H_0 : \mu \leq 60$, v.s. $H_a : \mu > 60$. $\alpha = .01$, $t = \dfrac{65 - 60}{15/\sqrt{100}} = 3.33$. Can view it as a z (assuming $s \approx \sigma$) & compare to $z_1^* = 2.33$. Can view it as a t and compare to t-table with 99 $d.f.$. In either case 3.33 is higher than the table value. So reject H_0 in favor of $\mu > 60$.

B. $H_0 : p \leq 1/2$, v.s. $H_a : p > 1/2$. $\alpha = .05$, $z = \dfrac{.59 - .5}{\sqrt{.5 \times .5}/\sqrt{100}} = 1.8 > z_1^* = 1.645$. Reject H_0 in favor of H_a: proportion of score 60 exceeds 50%.

C. $n = \dfrac{(z^*)^2(1/4)}{e^2} = \dfrac{(1.96)^2(1/4)}{.07^2} = 196$. If use Wilson can subtract 4 to get 192.

2. A. $H_0 : \mu_1 \geq \mu_2$ v.s. $H_a : \mu_1 < \mu_2$, $\alpha = .05$. Since sample size is large can assume $s_1 \approx \sigma_1$ and $s_2 \approx \sigma_2$ to get

$$z = \frac{65 - 70}{\sqrt{\frac{15^2}{100} + \frac{13^2}{100}}} = -2.52 < -1.645.$$

Reject H_0. Or one can use pooled t-test:

$$s^2_{\text{pooled}} = \frac{99}{198}(15)^2 + \frac{99}{198}(13)^2 = 197.$$

$$t = \frac{65 - 70}{\sqrt{197}\sqrt{\frac{1}{100} + \frac{1}{100}}} = -2.52$$

Compare to t-table and reject H_0.

B. $H_0 : p_1 \geq p_2$ v.s. $H_a : p_1 < p_2$, $\alpha = .05$.

$$\bar{p} = \frac{59 + 69}{200} = .64.$$

$$z = \frac{.59 - .69}{\sqrt{.64 \times .36}\sqrt{\frac{1}{100} + \frac{1}{100}}} = -1.473 > -1.645.$$

Don't reject H_0.

C. (i). a). Since μ_1 is closer to μ_2 in situation 1. So the probability to make a Type II error is higher in Situation 1.

(ii). c). Probability of making a type I error is placed at $\alpha = .05$.

(iii). c). p-value only depends on the null.

(iv). a). You need a larger sample size to make a .05 error when the means are closer than when they are further apart.

3. A. i). Percentage of the variability in 1993 returns that is explained by 1992 returns is 4.2%.

ii). If there were no relation between the two returns (R^2 =0), you would observe an R^2 of 4.2% or higher only .0399 (or about 4%) of the time.

B. i). $\widehat{y} = 16.65 - .3 \times \text{return } 1992 = 16.65 - .3 \times 10 = 13.65$.

ii).RMSE is 11.55 to get a range of $13.54 \pm 1.96 \times 11.55$ or 13.65 ± 22.54.

C. i). Data are clearly paired - Figure 3. $H_0 : \mu_d = 0, \text{v.s.} H_a : \mu_d \neq 0$. $t = \frac{\bar{x}}{s/\sqrt{n}} = \frac{7.2567}{15.549585}\sqrt{100} = 4.7$. Since $|t| > t^*$ (99 $d.f.$ t-value ≈ 2). Reject the null hypothesis that the mean returns are the same in favor of there is a difference. (could also use z-test).

ii). It is clear from the box-plot of diiference & normal quantile plot that the difference in returns are very skewed. Also, ther are a few very large differences of over 80.

D.

$$\text{var}\Big(\frac{1}{2}A + \frac{1}{2}B\Big) = \frac{1}{4}\text{var}(A) + \frac{1}{4}\text{var}(B) + 2 \cdot \frac{1}{2} \cdot \frac{1}{2} \cdot \sigma_A\sigma_B\text{Corr.}$$
$$= 61 + 60\text{Corr.} \leq var(A) = 100.$$

$$\Rightarrow 60\text{Corr.} \leq 39.$$

$$\Rightarrow \text{Corr.} \leq \frac{39}{60} = .65.$$

4. A. $H_0 : \mu \geq 4, \text{v.s.} H_a : \mu < 4. \alpha = .05, \sigma = 3$. $z = \frac{2.1 - 4}{3/\sqrt{5}} = -1.416 > -1.645$. Cannot reject H_0, so the vaccine should not be released based on this study.

B. Cut off $= \mu_0 - z_1^* \frac{\sigma}{\sqrt{n}} 4 - 1.645 \frac{3}{\sqrt{5}} = 1.79$.

$$P(\bar{x} < 1.79 | \mu = 2) = P\Big(z < \frac{1.79 - 2}{3/\sqrt{5}}\Big) = 0.4378.$$

C.

$$n = \frac{(z_\alpha + z_\beta)^2 \sigma^2}{(\mu_\alpha - \mu_0)^2} \frac{(1.645 + 1.28)^2 \times 3^2}{(2-4)^2} = 19.25 \text{ or } 20.$$

D.

Lot	z
1	$\frac{2.7-4}{3/\sqrt{18}} = -1.84$
2	$\frac{2.5-4}{3/\sqrt{8}} = -1.41$
1	$\frac{2.9-4}{3/\sqrt{16}} = -1.47$
1	$\frac{2.6-4}{3/\sqrt{11}} = -1.55$

So if Lot 3 is released, Lot 4 must also be released because it has a lower z-value than Lot 3.

5. A. i). Reject only with 4 stars.

ii). $P(4|H_a) = 0.38 \Rightarrow P(\{1,2,3\}|H_a) = .62$.

iii). $P(\geq 3|H_0) = .16 + .05 = .21$.

B. i).

$$P(\text{success}|3) = \frac{.25 \times .32}{.25 \times .32 + .75 \times .16} = .4.$$

ii). $E(\text{Profit}|3) = 50 \times P(\text{success}|3) + (-25) \times P(\text{bust}|3) = 50 \times .4 - 25 \times .6 = 5.$.

So on average the movies rated at 3 are profitable.