MATH 143

Inference practice problems

1. (a) 1-proportion confidence interval

   (b) We seek n (sample size) so that $0.03 \geq \text{(m.o.e.)} = (z \text{ critical val.})\sqrt{p(1-p)/n}$. With no estimate for p, we use the fact that $p(1-p)$ has a maximum value of 0.25 (when $p = 0.5$). Thus, it suffices to solve the inequality

   $$0.03 \geq 1.96\sqrt{\frac{0.25}{n}} \qquad \Rightarrow \qquad n \geq 1067.111.$$

   Since n must be a whole number, we need n to be at least 1068.

   (c) CI: $\hat{p} \pm 1.96\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = 0.2103 \pm 1.96\sqrt{\frac{(0.2103)(0.7897)}{252}}$, or $(0.160, 0.261)$.

   (d) The appropriate null/alt. hypotheses are

   H_0: $p = 0.85$, $\quad H_a$: $p < 0.85$,

   where the sample statistic for customer *satisfaction* is $\hat{p} = 199/252 \approx 0.7897$. So, the test z-statistic is

   $$z = \frac{0.7897 - 0.85}{0.02567145} \approx -2.35 \qquad \Rightarrow \qquad P(Z < -2.35) = 0.0094.$$

   Thus, we reject the null hypothesis, concluding that customer satisfaction is unacceptably low. $(z = -2.35, P = 0.0094)$

2. This data calls for a chi-square test, with hypotheses

   H_0: same proportion of crimes are committed each day of the week
   $p_{\text{Sun}} = p_{\text{Mon}} = \cdots = p_{\text{Sat}} = 1/7$.
   H_a: proportion of crimes is not the same for each day of the week.

   Under the null hypothesis, the expected value of crimes for each day would be 40 (that is, we would expect $p_{\text{Sun}} = p_{\text{Mon}} = \cdots = p_{\text{Sat}} = 1/7$). Thus, our test statistic is

   $$\chi^2 = \frac{1}{40}\left(8^2 + 2^2 + 9^2 + 0^2 + 1^2 + 0^2 + 4^2\right) = \frac{166}{40} = 4.15.$$

   Here the number of degrees of freedom is $df = 7 - 1 = 6$, so $P(\chi^2 > 4.15) > 0.10$ (actually, much greater than 0.10). Thus, at any reasonable significance level we find this sample data insufficiently different (as measured by the test statistic) from what we would have expected under H_0 to reject it. In other words, the sample data is consistent with the belief that crimes are committed with the same frequency regardless of the day of the week. $(\chi^2 = 4.15, df = 6, P > 0.10)$

3. (a) 1-sample t confidence interval

   (b) Even if these students from State Univ. were selected at random, the sampling frame is wrong (undercoverage) for drawing a conclusion about (all) college students. Perhaps a conclusion about students at State U. is appropriate.

4. (a) One might first think of the 2-sample t test (in which case 1-way ANOVA would also be a possibility). But the grades from the two classes are *not independent.* In fact, they are paired in a natural way. So, we should use a paired t-test.

   (b) $H_0 : \mu_{\text{diff}} = 0, \quad H_a : \mu_{\text{diff}} \neq 0.$

   (c) The test statistic is

$$t = \frac{0.17 - 0}{0.32/sqrt20} \approx 2.376.$$

   From Table D (with $df = 19$) we see that

$$P(|t| > 2.376) = 2P(t > 2.376) = 2(\text{number between 0.005 and 0.01}),$$

   yielding a P-value between 0.01 and 0.02. At the 1% significance level we do not see sufficient evidence to conclude grades are different between the two courses. ($t = 2.376, df = 19, 0.01 < P < 0.02$)

5. (a) $H_0 : \mu = 50, \quad H_a : \mu > 50.$

   (b) No. The full population distribution would include hours of operation between repairs for all such AC units. Nevertheless, a sample of 50 begins to reveal just how such a population distribution would appear.

   (c) No. Despite the fact that it gives us the impression the population distribution is quite skewed to the right, our sample size of 52 is adequate to insure that the sampling distribution for $\bar{x}$ is nearly normal.

6. (a) temperature (explanatory), stirring rate (explanatory) and yield (response)

   (b) 6

   (c) two-way ANOVA

   (d) 11

7. (a) Yes

   (b) 2-sample t procedures

   (c) Yes, at all three levels. In fact, we may say that the P-value of a test of significance is < 0.005.

   (d) 0.45, half way between the two extremes of the CI.

8. (a) increase, (b) decrease, (c) increase

9. (a) H_0 : The mean difference in ages is zero ($\mu_B - \mu_G = 0$).
   H_a : The mean difference in ages is nonzero ($\mu_B - \mu_G \neq 0$).

   (b) The relevant test statistic is

$$t = \frac{(12.2 - 11.9) - 0}{\sqrt{3.4^2/81 + 4^2/88}} = 0.526.$$

   Consulting Table D with $df = 80$, we see $P(|t| > 0.526) = 2P(t > 0.526) > (2)(0.1) = 0.2$. Thus, a sample whose t-statistic is as extreme (as far or farther from zero) as ours occurs more than 20% of the time when H_0 is true. We should not reject H_0.

(c) The value $12.2 - 11.9 = 0.3$ will be at the center of the CI, and zero will be inside the interval as well. The actual 95% CI is

$$0.3 \pm (1.99)\sqrt{\frac{3.4^2}{81} + \frac{4^2}{88}}, \qquad \text{or} \qquad (-0.834, 1.434).$$

10. (a) 1-way ANOVA

(b) $df_1 = 3$, $df_2 = 74$

(c) $P < 0.001$

(d) Each individual test is a 2-sample t test. We have to adjust the significance level (something that is done for us automatically using the Bonferroni procedure) because performing multiple tests at the 5% level increases our odds (above 5%) of rejecting at least one null hypothesis when it is true.

(e) There is a significant difference in the amount of improvement one sees in the "piano" children and those in all other groups (i.e., piano vs. computer, piano vs. singing and piano vs. none). No other pairwise comparisons show significant differences. The "piano" children have done the best for all groups. (Note: This is a statement about population means, not just sample means.) Ranking performance in the sample means, we have "singing" then "computer" then "none" then "piano".

11. (a) H_0: The proportion of responders is the same regardless of whether a prenotification letter is sent ($p_1 - p_2 = 0$).

H_a: The prenotification letter plays a role in whether or not we get a survey response ($p_1 - p_2 \neq 0$).

Both chi-squared and a 2-proportion test are appropriate. Note that the expected cell counts for the associated two-way table are all much higher than 5. (A two-way table with the *actual* counts is found on the top of p. 646.) The 2-proportion test is applicable because the two-way table is (2 rows)-by-(2 cols).

(b) **The 2-proportion test**:

In what follows, the subscript '1' refers to the population receiving the prenotification letter. Then

$$\hat{p}_1 = \frac{2570}{5018} \approx 0.5122 \qquad \text{and} \qquad \hat{p}_2 = \frac{2645}{5029} \approx 0.5259.$$

For 2-proportion tests of significance, we used the pooled proportion estimate in the calculation of SE:

$$\hat{p} = \frac{2570 + 2645}{5018 + 5029} \approx 0.5191, \qquad \text{so} \qquad \text{SE} = \sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)} \approx 0.009969307.$$

This yields a test z-statistic

$$z = \frac{(0.5122 - 0.5259) - 0}{0.009969307} \approx -1.374.$$

The P-value is

$$P(|Z| > 1.374) = 2P(Z < -1.374) \approx (2)(0.0853) = 0.1706.$$

Thus, there is insufficient evidence in this sample to reject H_0—that is, the null hypothesis is consistent with our sample data.

The chi-squared test:

The two-way table with expected counts in parentheses is

	Response		
	Yes	No	Total
Letter	2570 (2604.645)	2448 (2413.355)	5018
No Letter	2645 (2610.355)	2384 (2418.645)	5029
Total	5215	4832	10047

We get test statistic

$$\begin{aligned} \chi^2 &= \frac{(2570 - 2604.645)^2}{2604.645} + \frac{(2448 - 2413.355)^2}{2413.355} + \frac{(2645 - 2610.355)^2}{2610.355} + \frac{(2384 - 2418.645)^2}{2418.645} \\ &\approx 1.9142. \end{aligned}$$

Checking Table F with $df = 1$, we get $P(\chi^2 > 1.9142)$ is between 0.15 and 0.2. We draw the same conclusion as we did from the 2-proportion test.