POS 3713 Summer 2000
Solutions to Review Problems for Midterm Exam #2

Note: Xbar denotes the sample mean

1. H_o: $\mu = 39.7$
   H_1: $\mu \neq 39.7$

We should establish the critical region using the Student t distribution because the population standard deviation (σ) is unknown, and N < 100. The df = N - 1 = 12 - 1 = 11. The test is two-tailed, and we can see in Appendix B that the critical two-tailed t-value for $\alpha = .05$ is +/- 2.201.

Now, we need to calculate $t = \dfrac{\text{Xbar} - \mu}{S/\sqrt{(N-1)}} = \dfrac{50.4 - 39.7}{11.8/\sqrt{(12-1)}} = 10.7/3.558 = 3.01$

Since our calculated (or obtained) value of t is greater than the critical value of t, we reject the null hypothesis and conclude that the jury is different from Leon county's population with respect to age.

2. H_o: $\mu = 440$
   H_1: $\mu > 440$

We use Z because N is greater than 100. This is a one-tailed test, thus we need to find the value of Z that has an area beyond Z of .05. If we look in Appendix A, we will see that this value falls between 1.64 and 1.65, so we will split the difference and use 1.645 as our critical value of Z.

$Z = \dfrac{\text{Xbar} - \mu}{s/\sqrt{N-1}} = \dfrac{458 - 440}{20/\sqrt{(126-1)}} = 18/1.79 = 10.06$

Since 10.06 > 1.645, we reject H_o and conclude that the local average is significantly larger than the national average.

3.

a. Since N < 100, we use a two-tailed t critical value for df = N - 1 = 85 - 1 = 84. The closest value is for df=60 (t=2.00) or df=120 (t=1.98).

c.i. = Xbar +/- $t(s/\sqrt{N-1})$ = 67.26 +/- $2.00(28.61/\sqrt{(85-1)})$ = 67.26 +/- 6.24

Interpretation: We are 95% confident that the mean percentage of females who read in the population is between 61% and 73.5%.

b. Yes, there is a significant difference in male and female literacy rates. The mean percentage of males who read is 11.47% above the mean percentage of females who read. The median is much higher for men (87 versus 71), while the mode is the same (100). The dispersion is also larger for females (variance is 818.38 for women compared to 418.01 for men, and standard deviation is 28.61 for women and 20.45 for men), which we can also see by the larger range for females (91 versus 72). Thus on average, men have a higher literacy rate than women and the variation in literacy rates across countries is much larger with respect to women. This can be explained by many factors. In poorer, more traditional countries, women serve a greater role in the household, and thus may not receive the same level of education as men.

4. 1963: $P_s = 160/200 = .8$; $\pi = P_s \pm Z_{\alpha/2}\cdot\sqrt{[P_u(1-P_u)]/N} = .8 \pm 1.96\cdot\sqrt{[.5(1-.5)]/200} = .8 \pm .069$

Interpretation: We are 95% confident that 73.1% to 86.9% of Americans approved of Johnson in 1963.

1968: $P_s = 70/200 = .35$; $\pi = P_s \pm Z_{\alpha/2}\cdot\sqrt{[P_u(1-P_u)]/N} = .35 \pm 1.96\cdot\sqrt{[.5(1-.5)]/200} = .35 \pm .069$

Interpretation: We are 95% confident that 28.1% to 41.9% of Americans approved of Johnson in 1968.

5. $H_o: P_u = .6$ N = 120
$H_1: P_u > .6$

$P_s = 84/120 = .7$
Since N > 100, we must calculate a Z score. This is a one tailed test, thus the critical Z value is 1.645.

$$Z = \frac{P_s - P_u}{\sqrt{[(P_u(1-P_u))/N]}} = \frac{.7 - .6}{\sqrt{[(.6)(.4)/120]}} = .1/.0447 = 2.24$$

Since the calculated Z of 2.24 exceeds the critical Z of 1.645, we reject the null hypothesis and conclude that drivers in Florida are more likely to speed. Governor Bush should be concerned about these results.

6. We calculate the IQV as follows: $IQV = \frac{k\,(N^2 - \Sigma f^2)}{N^2\,(k - 1)}$

$$IQV = \frac{3(1669^2 - (612^2 + 513^2 + 544^2))}{1669^2\,(3-1)} = \frac{3(2785561 - 933649)}{5571122} = \frac{5555736}{5571122}$$

IQV = 0.997

Interpretation: The IQV is extremely close to 1, indicating that party identification is quite dispersed. In other words, each party is likely to attract approximately the same number of people (recall that IQV = 1 when the frequencies in each category are equal).

6.

Descriptive Statistics

	N	Range	Minimum	Maximum	Mean	Std. Deviation	Variance
Clinton thermometer	1705	100	0	100	59.34	29.58	874.802
Hillary Clinton Thermometer	1685	100	0	100	52.81	29.85	891.005
Valid N (listwise)	1683						

a) $Z = \frac{X - Xbar}{s} = \frac{75 - 59.34}{29.58} = \frac{15.66}{29.58} = 0.53$, Pr(Clinther > 75) = Pr(Z > 0.53) = .2981

$Z = \frac{X - Xbar}{s} = \frac{75 - 52.81}{29.85} = \frac{22.19}{29.85} = 0.74$, Pr(Hillther > 75) = Pr(Z > 0.74) = .2297

b) $Z = \frac{X - Xbar}{s} = \frac{25 - 59.34}{29.58} = \frac{-34.34}{29.58} = -1.16$, Pr(Clinther < 25) = Pr(Z < -1.16) = 0.123

$Z = \frac{X - Xbar}{s} = \frac{25 - 52.81}{29.85} = \frac{-27.81}{29.85} = -0.93$, Pr(Hillther < 25) = Pr(Z < -0.93) = .1762

c) $Z = \frac{X - Xbar}{s} = \frac{30 - 59.34}{29.58} = \frac{-29.34}{29.58} = -0.99$

$Z = \frac{X - Xbar}{s} = \frac{70 - 59.34}{29.58} = \frac{10.66}{29.58} = 0.36$

Pr(30 < Clinther < 70) = Pr(Z > -0.99 and Z < 0.36) = .3389 + .1406 = .4795 or 47.95%

$Z = \frac{X - Xbar}{s} = \frac{30 - 52.81}{29.85} = \frac{-22.81}{29.85} = -0.76$

$Z = \frac{X - Xbar}{s} = \frac{70 - 52.81}{29.85} = \frac{17.19}{29.85} = 0.58$

Pr(30 < Hillther < 70) = Pr(Z > -0.76 and Z < 0.58) = .2764 + .2190 = .4954 or 49.54%

d) $Z = \frac{X - Xbar}{s} = \frac{85 - 59.34}{29.58} = \frac{25.66}{29.58} = 0.87$, Pr(Clinther < 85) = Pr(Z < 0.87) = .8087

A score of 85 on the Bill Clinton feeling thermometer is in the 80.87th percentile.

$Z = \frac{X - Xbar}{s} = \frac{85 - 52.81}{29.85} = \frac{32.19}{29.85} = 1.08$, Pr(Hillther < 85) = Pr(Z < 1.08) = .8599

A score of 85 on the Hillary Clinton feeling thermometer is in the 85.99th percentile.

7. Mean = (70 + 72 + 68 + 65 + 71 + 75 + 66 + 67 + 65 + 70)/10 = 689/10 = 68.9
Median: arrange values in ascending order

65 65 66 67 68 70 70 71 72 75

Since N is even, we take the average of the 2 middle values: (68 + 70)/2 = 69
Mode: most frequently occurring value; in this case we have 2 modes: 65 and 70.
Range = highest value - lowest value = 75 - 65 = 10

Variance = $s^2 = \frac{\Sigma (X_i - Xbar)^2}{N - 1}$

$s^2 = [(70 - 68.9)^2 + (72 - 68.9)^2 +(68 - 68.9)^2 +(65 - 68.9)^2 +(71 - 68.9)^2 +(75 - 68.9)^2 +(66 - 68.9)^2 +(67 - 68.9)^2 +(65 - 68.9)^2 +(70 - 68.9)^2]/(10 - 1)$

$s^2 = 10.77$

Standard deviation = $s = \sqrt{s^2} = \sqrt{10.77} = 3.28$

Interpretation: On average, this golfer scores 3.28 points above or below his/her mean score of 68.9.

8.

a. What is the independent variable for this theory? What is the dependent variable?

IV: sex, DV: party identification

b. What *percentage* of the total sample is male?

47.1%

c. What *proportion* of the total sample is Democratic?

612/1669 = .367

d. What *percentage* of respondents are female Independents?

244/1669 = .146 or 14.6%

e. How many females in this sample identify with the Republican party (i.e., what is the frequency)?

271

f. Does the table provide any evidence to support the theoretical hypothesis? Justify your answer using the information presented.

Yes, the table does provide evidence to support the hypothesis that women are more likely to identify with the Democratic party. The percentage of women identifying Democratic is 41.7%, which is much higher than the percentage of women identifying with the Republican party (30.7%) or declaring themselves Independent (27.6%).