Math 220 : Elementary Statistics — Kane Nashimoto
Final Exam Solutions
Fall 2007

1. Graphical and numerical summaries of data.

   (a)

Stem	Leaf
5	0
6	00
7	055
8	05
9	0
10	
11	0

   The distribution of the data is positively skewed.

   (b) Note : $\sum x = 75.5$; $\sum x^2 = 596.75$; $n = 10$

   $$\bar{x} = \frac{\sum x}{n} = \frac{75.5}{10} = 7.550$$

   $$s = +\sqrt{\frac{\sum x^2 - (\sum x)^2/n}{n-1}} = +\sqrt{\frac{596.75 - (75.5)^2/10}{10-1}} = 1.723$$

2. Bias in scientific studies.

   Non-response bias - an example of a study in which responses are not obtained from all participants.

3. Inference about proportion.

   (a) $(.33)(.31) = .102 \;\Rightarrow\; 10.2\%$

   (b) Use $\hat{p} = .102$.

   $$n = \frac{\hat{p}(1-\hat{p})\, z^2_{.10/2}}{m^2} = \frac{.102(1-.102)(1.645)^2}{(.01)^2} = 2478.61 \;\Rightarrow\; 2479$$

4. Binomial random variable.

   Note : X, num of people willing to buy a GM car, is binomial with $n = 15$ and $p = .25$.

   (a) $P(X \geq 5) = 1 - P(X \leq 4) = 1 - .6865 = .3135$

   (b) $P(3 \leq X \leq 6) = P(X \leq 6) - P(X \leq 2) = .9434 - .2361 = .7073$

5. Normal random variable.

   Note : X, heating bill in December, is normal with $\mu = 150$ and $\sigma = 25$.

   (a) $P(X < 120) = P\left(\frac{X-\mu}{\sigma} < \frac{120-150}{25}\right) = P(Z < -1.20) = .1151$

   (b) $z = 1.28 \;\because\; P(Z \leq 1.28) \approx .90$

   $$z = \frac{x-\mu}{\sigma}\; ;\; 1.28 = \frac{x-150}{25}\; ;\; x = 150 + (1.28)(25) = 182.00$$

6. Probability.

   A: {electrical problem causing breakdown}
   B: {mechanical problem causing breakdown}
   Note: $P(A) = .30$; $P(B) = .60$; $P(A \cap B) = .15$

   (a) $P(A \cup B) = P(A) + P(B) - P(A \cap B) = .30 + .60 - .15 = .75$
   (b) $P(A \cap B) = .15$ but $P(A)P(B) = (.30)(.60) = .18$
   $\Rightarrow$ Two events are not independent.

7. Chi-square test of homogeneity.

   (a) H_0: Relative prop of D and R supporters are homogeneous in the two cities.
   H_a: Relative prop of D and R supporters are not homogeneous in the two cities.

   Observed and marginal counts:

47	31	78
50	67	117
97	98	195

   $\widehat{n_{ij}} = (n_{i\cdot})(n_{\cdot j})/n$
   $\widehat{n_{11}} = (78)(97)/195 = 38.80$ $\widehat{n_{12}} = (78)(98)/195 = 39.20$
   $\widehat{n_{21}} = (117)(97)/195 = 58.20$ $\widehat{n_{22}} = (117)(98)/195 = 58.80$

   $$h^* = \sum_i \sum_j \frac{(n_{ij} - \widehat{n_{ij}})^2}{\widehat{n_{ij}}}$$
   $$= \frac{(47-38.80)^2}{38.80} + \frac{(31-39.20)^2}{39.20} + \frac{(50-58.20)^2}{58.20} + \frac{(67-58.80)^2}{58.80} = 5.747$$

   $h_{.025,\,(2-1)(2-1)} = 5.024 < h^* = 5.747 < h_{.01,\,(2-1)(2-1)} = 6.635$
   $.01 < p\text{-value} < .025$
   Reject H_0 at $\alpha = .05$.
   (There are more supporters of a Democratic candidate in the city in Minnesota, and there are more supporters of a Republican candidate in the city in Texas.)

   (b) $\widehat{n_{ij}} > 5$ for all i, j
   $\Rightarrow$ Assumption is satisfied.

8. Inference about mean.

   (a) $\bar{x} \pm t_{.05/2,\,30-1}\dfrac{s}{\sqrt{n}}$; $523 \pm 2.045\dfrac{116}{\sqrt{30}}$; 523 ± 43.31; $(479.69,\ 566.31)$
   (b) $H_0\colon \mu = 500$ vs. $H_a\colon \mu \neq 500$

   $$t^* = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} = \frac{523 - 500}{116/\sqrt{30}} = 1.086$$

   $t_{.15,\,30-1} = 1.055 < |t^*| = 1.086 < t_{.10,\,30-1} = 1.311$
   $.20 < p\text{-value} < .30$
   Retain H_0 at $\alpha = .10$. ($\mu \approx 500$)

9. Paired-samples t test by SPSS.

   $H_0\colon \mu_\delta = 0$ vs. $H_a\colon \mu_\delta < 0$ ("follow-up" minus "baseline")
   $t^* = -2.584$
   one-sided p-value $= .019 \div 2 = .0095$
   Reject H_0 at $\alpha = .05$.
   The mean body weight has declined significantly after the weight-loss program.