Statistics 515 - Solutions for Spring 2003 Practice Exam

Part I:

1) Mississippi has a population of approximately 3,000,000 and is represented by a square 2cm by 2cm. How long will each side of the square for Pennsylvania (population 12,000,000) be? **Area needs to be four times larger than 4 cm^2 = 16cm^2, so sqrt(16)cm x sqrt(16)cm or 4cm x 4 cm**

2) A data-entry employee is entering a large list of salaries and one of them is mistyped by either adding or deleting 0s from the end. Is this mistake more likely to affect the mean or the median of the salaries? **Mean**

3) If the data is approximately normal (or bell-shaped), about what percent of the data will be considered possible outliers?
1-.997 = .003 =0.3%
If the data is skewed, what is the largest percent of the data that could be considered possible outliers?
$1 - (1-1/k^2) = 1/k^2 = 1/3^2 = 1/9 \approx 11.1\%$

4) P(A)=0.6, P(B)=0.4. If A and B are mutually exclusive, what is $P(A\cap B)$? **0**

5) P(A)=0.6, P(B)=0.4. If P(A|B)=0.5, what is P(AUB)? **$P(AUB)=P(A)+P(B)-P(A\cap B)= P(A)+P(B)-P(B)P(A|B)$**
=0.6+0.4-(0.4)(0.5)=0.8

6) Let the random variable X have the following distribution:

x	0	2	6
p(x)	0.5	0.25	0.25

mean = 0(0.5) + 2(0.25) + 6(0.25) = 2 **variance = $(0-2)^2(0.5) + (2-2)^2(0.25) + (6-2)^2(0.25) = 2+0+4=6$**

7) X is a normal random variable with μ=25, σ^2=25, and σ=5. Find $P(X \leq 30)$. **$= P((X-\mu)/\sigma \leq (30-25)/5)$**
$= P(Z \leq 1) = 0.5+0.3413=0.8413$

8) Z is a standard normal random variable. Find z_0 such that $P(Z \geq z_0) = 0.0778$. **Look up 0.4222 on the table to get 1.42**

Part II:

1a) Find the mean. **(10cm+9cm+10cm+17cm+25cm+19cm)/6 = 90cm/6 = 15cm**

b) Find the median. 9cm 10cm 10cm 17cm 19cm 25cm **The average of 10cm and 17cm is 13.5cm.**

c) Find the mode **10cm occurs most often**

d) Find the variance. **$((10cm-15cm)^2+ (9cm-15cm)^2 + (10cm-15cm)^2+ (17cm-15cm)^2 + (25cm-15cm)^2+(19cm-15cm)^2)/(6-1)$**
$= (25\ cm^2 +36cm^2 +25cm^2 +4cm^2+100cm^2+16cm^2)/5 = 206cm^2/5 = 41.2\ cm^2$

e) Find the standard deviation. **sqrt(41.2 cm^2) ≈ 6.419 cm**

2) Based on past censuses and computer models, it is predicted that a large population is 51% female. A survey of 3,000 randomly chosen residents is conducted. Because the population is large and the residents are randomly chosen, this could be considered a binomial experiment.

a) What is the expected number of females in the 3,000 people surveyed? **μ=np=3,000(0.51)=1530**

b) According to the past information, what is the standard deviation for the number of females in the 3,000 people surveyed?
$\sigma=\sqrt{np(1-p)}=\sqrt{3,000(0.51)(1-0.51)}=\sqrt{3,000(0.51)(0.49)}=\sqrt{749.7}\approx 27.38$

c) According to the past information, what is the probability that exactly 1,500 of those surveyed will be women? (You do not need to simplify your answer). **(3000 choose 1500) $(0.51)^{1500}$ $(0.49)^{3000-1500}$**

d) Assume that the number of women found by doing this experiment is approximately normally distributed with the mean you found in a and the standard deviation you found in b. Approximately what is the probability that you will observe 1,500 or fewer females? That is, find the approximate value of $P(X\leq 1500)$. **$P(X\leq 1500) = P((X-\mu)/\sigma \leq (1500-1530)/ 27.38) = P(Z \leq -1.10)$**
= 0.5 – 0.3643 = 0.1357

e) One way to check whether we believe a binomial random variable should be approximately normally distributed would be to program a computer to repeat the experiment a large number of times. Say we programmed the computer to do this experiment 500 times. What kind of plot would you use to check that the results follow a normal distribution, and what would you look for in that plot? **Q-Q plot should look like a straight line.**