Stat/For/Hort 571 — Midterm I, Fall 2000
Solutions

1. (a) The stem and leaf display is useful for a small data set such as this.

```
6 | 1 4
6 | 6 7 7 8
7 | 0 2 4
7 | 9
```

The display appears roughly symmetrical, centered around 6.7. Median $y_{[0.50]} = 6.75$.

(b) Since $\bar{Y} \sim N(6.45, 0.30/15)$, $P(\bar{Y} < 6.60) = P(Z < \frac{6.60-6.45}{\sqrt{0.3/15}}) = P(Z < 1.06) = 0.8554$.

2. (a) $P(Y_1 \geq 7) = P(Y_1 = 7) + P(Y_1 = 8) = \frac{8!}{7!1!}0.7^7 0.3^1 + \frac{8!}{8!0!}0.7^8 0.3^0 = 0.1976 + 0.0576 = 0.2552$.

(b) $P(Y_2 \geq 7) = P(Z \geq \frac{7-5.6}{\sqrt{1.68}}) = P(Z \geq 1.08) = 0.1401$.

(c) $E(Y_1) = 8 \times 0.7 = 5.6 = E(Y_2)$ and $Var(Y_1) = 8 \times 0.7 \times 0.3 = 1.68 = Var(Y_2)$.

(d) Sample size is small with $n(1-p) = 2.4 < 5$; thus normal is not a good approximation to binomial. Also in approximating a discrete distribution with a continuous distribution, there will almost always be some difference between them.

3. (a) Since $n = 9$, the df are 8. If $V^2 \sim \chi^2_8$, then from tables $P(V^2 \geq 20.09) = 0.01$. But then we have

$$0.01 = P(S^2 > c) = P(V^2 > \frac{8}{25}c)$$

It follows that $8c/25 = 20.09$ and so $c = 62.78$.

(b) By independence, $P(S_1^2 > c \text{ and } S_2^2 > c) = P(S_1^2 > c) \times P(S_2^2 > c) = 0.01 \times 0.01 = 0.0001$.

4. (a) Let Y = amount of money you receive. $E(Y) = 0.80 \times 0 + 0.18 \times 1 + 0.02 \times 100 = 2.18$. Since this is greater than the cost of playing the game, 2, the company will lose money in the long run.

(b) $Var(Y) = 0.80 \times (0-2.18)^2 + 0.18 \times (1-2.18)^2 + 0.02 \times (100-2.18)^2 = 195.42 = 13.97^2$. Hence $\bar{Y}_{\text{NA}} \sim N(2.18, 195.42/100)$ by the central limit theorem (CLT) and $P(\bar{Y} \geq 2) \approx P(\bar{Y}_{\text{NA}} \geq 2) = P(Z \geq \frac{2-2.18}{1.397}) = P(Z \geq -0.129) = 0.5517$.

(c) The assumptions of a random sample are implicit. Also a large sample size n is needed for the CLT to provide a good approximation. Note that the distribution of Y is strongly skewed. Hence, according to our simulation, the distribution of $\bar{Y}$ is still somewhat skewed.

5. $E(\hat{p}) = p = 0.4$ and $Var(\hat{p}) = \frac{p(1-p)}{n} = \frac{0.24}{n}$. Hence, we have $\hat{p}_{\text{NA}} \sim N(0.4, \frac{0.24}{n})$ and

$$0.05 = P(\hat{p} > 0.45) \approx P(\hat{p}_{\text{NA}} > 0.45) = P(Z > \frac{0.45 - 0.4}{\sqrt{0.24/n}}).$$

From the tables, $P(Z > 1.645) = 0.05$. Thus $\frac{0.45-0.4}{\sqrt{0.24/n}} = 1.645$. It follows that $n = 259.77$ and is rounded up to 260. Normal approximation is justified, because $np > 5$ and $n(1-p) > 5$ with $n = 260$ and $p = 0.4$.

Grade Distribution

```
  100:17
90-99:68
80-89:40
70-79:19          mean = 86, median = 90
60-69:9
50-59:5
  <50:4
```