Stat/For/Hort 571 — Midterm I, Fall 98 — Brief Solutions

1. (a) One helpful display is the stem and leaf display. For this problem it is good to use the first digit as the stem, but stretch the stem out by putting numbers with "leaves" that are < 50 with one stem and those ≥ 50 with another stem, as in the following:

```
3 | 35
3 | 64, 84, 99
4 | 23, 11, 18, 02, 30
4 | 83, 77, 91, 59, 68, 76
5 | 04, 20, 43
5 | 61
```

The display is roughly symmetric and the center is around 450cm.

(b) Let μ be the mean height of this variety of seedlings. Thus we have

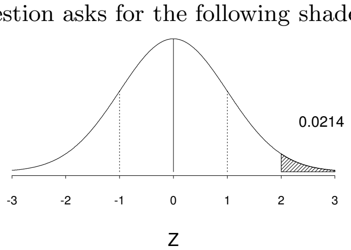
$$H_0 : \mu = 420 \text{ cm} \quad \text{and} \quad H_A : \mu \neq 420 \text{ cm}$$

We calculate $\bar{y} = 8548/19 \doteq 449.89$ and we are given $\sigma^2 = 3800$. So

$$z = \frac{\bar{y} - \mu}{\sigma/\sqrt{n}} = \frac{449.89 - 420}{\sqrt{3800}/\sqrt{19}} \doteq 2.11.$$

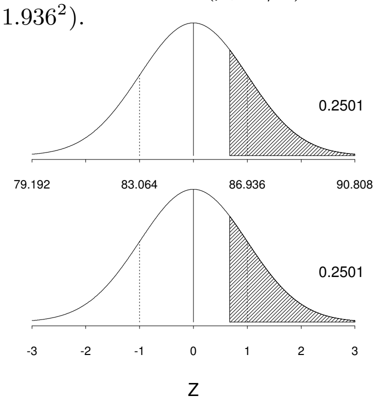
The p-value is $2P(Z \geq 2.11) = 2(0.0174) = 0.0348$. There is moderate evidence against H_0, and we reject the null at 10% and 5% since the p-value falls below these levels. However, we do not reject H_0 at 1%.

2. (a) The question asks for the following shaded area:



From the tables, we find that $P(Z \geq 2) = 0.0228$ and $P(Z \geq 3) = 0.0013$. Thus $P(2 \leq Z \leq 3) = 0.0228 - 0.0013 = 0.0215$. The desired area is twice this and is 0.043.

(b) We have that $\bar{W} \sim N(\mu, \sigma^2/n) = N(85, 60/16) \doteq N(85, 1.936^2)$.



The desired w_* is 0.67 standard deviations away from the mean since $P(Z \geq 0.67) = 0.25$, and we solve

$$\frac{w_* - 85}{1.936} = 0.67$$

so that $w_* = 85 + (0.67)(1.936) \doteq 86.3$.

3. (a) Since $Y \sim B(8, 0.3)$, we have that

$$\begin{aligned} & P(Y = 0) + P(Y = 1) \\ = \; & \frac{8!}{0!8!} 0.3^0 0.7^8 + \frac{8!}{1!7!} 0.3^1 0.7^7 \\ \doteq \; & 0.0576 + 0.198 \doteq 0.256. \end{aligned}$$

(b) Here $Y \sim B(80, 0.3)$. We appeal to the normal approximation, noting that $np = 24$ and $n(1-p) = 56$ so that this is appropriate. Thus $Y_{NA} \sim N(np, np(1-p)) \doteq N(24, 4.099^2)$, and the desired probability is $P(Y \leq 10) \doteq P(Y_{NA} \leq 10) = P(Z \leq (10-24)/4.099) = P(Z \leq -3.41) = 0.0003$.

(c) When n is large, the proportion Y/n becomes closer to p (with a standard deviation equal to $\sqrt{p(1-p)/n}$. Thus, it is much less likely to get a proportion that is less than or equal to 0.125 ($= 1/8 = 10/80$).

4. (a) $E(Y) = \mu = \sum yp(y) = 0 \times 0.2 + 1 \times 0.5 + 5 \times 0.3 = 2$ and $\text{Var}(Y) = \sigma^2 = \sum (y - \mu)^2 p(y) = (0-2)^2 \times 0.2 + (1-2)^2 \times 0.5 + (5-2)^2 \times 0.3 = 4$.

(b) By the Central Limit Theorem (CLT) (see part (c)), we have that $\bar{Y}$ is approximately $N(\mu, \sigma^2/n) = N(2, 0.2^2)$. Thus $P(\bar{Y} < 1.5) \doteq P(Z < (1.5-2)/0.2) = P(Z < -2.5) = P(Z > 2.5) = 0.0062$.

(c) To calculate the probability in (b) we need to appeal to the CLT which says that $\bar{Y}$ for a random sample is approximately normal provided that the sample size n is large. How large n must be depends on the population from which the observations are sampled. Typically the approximation is reasonable if $n \geq 30$ as long as the underlying population is not too skewed. The distribution here is not too heavily skewed.

5. (a) Since $n = 4$, $E(V^2) = n - 1 = 3$ and $\text{Var}(V^2) = 2(n-1) = 6$ so that the standard deviation $= \sqrt{6} \doteq 2.45$.

(b) As the mean is 3 and the standard deviation is 2.45, we are interested in the chance that V^2 falls within $3 \pm 2.45 = [0.55, 5.45]$, or $P(0.55 \leq V^2 \leq 5.45)$, as illustrated in the picture below:

Using Table B, we can bound separately the two parts that are not shaded, noting that df $= n - 1 = 3$. For the right side $0.1 = P(V^2 \geq 6.25) < P(V^2 \geq 5.45) < P(V^2 \geq 4.11) = 0.25$. Then for the left, $0.05 = 1 - 0.95 = P(V^2 \leq 0.35) < P(V^2 \leq 0.55) < P(V^2 \leq 0.58) = 1 - 0.9 = 0.1$.

The two nonshaded areas are then bounded between $0.1 + 0.05 = 0.15$ and $0.25 + 0.1 = 0.35$. Thus, the shaded area is bounded between $1 - 0.35 = 0.65$ and $1 - 0.15 = 0.85$. Note that because Table B is sparse, interpolation is dangerous. [Note an exact calculation with Minitab gives 0.766.]

Grade Distribution

```
  100:5
 95-99:34
 90-94:27
 80-89:26          n = 151
 70-79:29          mean = 82.98, median = 87
 60-69:21          quartiles = 73, 95
 50-59:6           s = 13.9
   <50:3
```