# Stat/For/Hort 571
# Midterm I, Fall 2004
# Brief Solutions

1. (a) A stem-and-leaf display:

```
1 | 7.1
1 | 8.6
2 | 0.4 0.9 1.8
2 | 2.5 3.6 3.7
2 | 4.2 5.3
2 | 6.4
2 | 8.0
```

The distribution appears to be fairly symmetric and unimodal centered around about 23. The standard deviation is $s = 3.16$.

(b) Let $X$ be the nitrogen concentration of a random plant. We have $X \sim N(27, 12.4)$.

$$\begin{aligned} \mathbb{P}(X < 25) &= \mathbb{P}\left(\frac{X-27}{\sqrt{12.4}} < \frac{25-27}{\sqrt{12.4}}\right) \\ &= \mathbb{P}(Z < -0.57) \\ &\doteq 0.2843 \end{aligned}$$

(c) The upper endpoint of the middle 99% is the 0.9950 quantile. The corresponding $z$-score is about $z = 2.575$. By symmetry, the lower endpoint has $z = -2.575$. The distance from the mean is $(2.575)(\sqrt{12.4}) \doteq 9.1$, so the end points are 17.9 and 36.1.

2. (a) Let $Y$ be the number of successful insertions. We model $Y \sim \mathrm{B}(9, 0.1)$.

$$\begin{aligned} \mathbb{P}(Y \geq 2) &= 1 - \mathbb{P}(Y=0) - \mathbb{P}(Y=1) \\ &= 1 - \frac{9!}{0!9!}(0.1)^0(0.9)^9 \\ &\quad - \frac{9!}{1!8!}(0.1)^1(0.9)^8 \\ &\doteq 1 - 0.3874 - 0.3874 = 0.2252. \end{aligned}$$

(b) $\mathbb{P}(Y = 9) = \frac{9!}{9!0!}(0.1)^9(0.9)^0 = 10^{-9}$

(c) Gene insertions are successful with low probability, 0.1. In nine independent trials, the mean number of successes is less than one, so the outcomes of zero and one success will be the most probable. The probability of two or more successes will certainly be much higher than the probability of nine successes whose probability is nearly zero.

(d) The mean is $\mu = 60(0.1) = 6$ and the variance is $\sigma^2 = 60(0.1)(0.9) = 5.4$.

3. (a) Events $W$ and $F$ are not disjoint because $\mathbb{P}(W \text{ AND } F) = 0.4 > 0$. That is, both W and F can co-occur with non-zero probability.

(b) Events $W$ and $F$ are not independent because $\mathbb{P}(W)\,\mathbb{P}(F) = (0.6)(0.8) = 0.48 \neq 0.4$.

(c) There are four possible states for the cats. $\mathbb{P}(W \text{ AND } F) = 0.4$, $\mathbb{P}(W \text{ AND NOT } F) = 0.6 - 0.4 = 0.2$, $\mathbb{P}(\text{NOT } W \text{ AND } F) = 0.8 - 0.4 = 0.4$, leaving $\mathbb{P}(\text{NOT } W \text{ AND NOT } F) = 0.0$ as these disjoint probabilities must sum to one. If $Y$ is the number of cats that climb the tree, then $\mathbb{P}(Y = 0) = \mathbb{P}(\text{NOT } W \text{ AND NOT } F) = 0.0$.

(d) From part(c), we can find $\mathbb{P}(Y = 1) = 0.6$ and $\mathbb{P}(Y = 2) = 0.4$. Thus, using the definition of expected value, $\mathbb{E}(Y) = (1 \times 0.6) + (2 \times 0.4) = 1.4$ Alternatively, if we let $I_W$ and $I_F$ be "1" or "0" if Wanda and Frieda are (or are not) in the tree, respectively, it follows that $Y = I_W + I_F$. Expected values of 0/1 random variables are the probability that they are one and, by the rule for expectation of a sum, $\mathbb{E}(Y) = \mathbb{E}(I_W) + \mathbb{E}(I_F) = 0.6 + 0.8 = 1.4$.

4. If $W_i$ is the winnings and $G_i = W_i - 15$ is the gain on the $i$th play, we have that $\mathbb{E}(G_i) = \mathbb{E}(W_i) - 15 = 0(0.9) + 50(0.09) + 1000(0.01) - 15 = -0.5$. We also can find that $\mathrm{Var}(G_i) = \mathrm{Var}(W_i) = (14.5)^2(0.9) + (35.5)^2(0.09) + (985.5)^2(0.01) = 10,014.75$.

(a) Our only hope is the central limit theorem. We let $\bar{G}_{NA} \sim N(-0.5, 10014.75/150)$.

$$\begin{aligned} \mathbb{P}(\bar{G} \geq 0) &\approx \mathbb{P}(\bar{G}_{NA} \geq 0) \\ &= \mathbb{P}\left(Z \geq \frac{0 - (-0.5)}{\sqrt{10014.75/150}}\right) \\ &= \mathbb{P}(Z \geq 0.06) \doteq 0.4761. \end{aligned}$$

(b) The gain on any single play is either $-15$, 35, or 985. It is impossible for any $G_i$ to be less than $-15$ and, thus, the probability of losing more than \$15 on average is zero.

(c) For problem (a), we assumed that the sampling distribution of the sample mean was approximately normal. For the highly skewed distribution, even a sample of size 150 does not guarantee that this approximation is good. Had we used the approximation for (b), we would have obtained a probability of 0.037 whereas the probability must be "0.0" from the reality of the problem..

### Grade Distribution

```
  100: 4
90-99:53
80-89:50
70-79:14          mean = 84.4, median = 88
60-69:11
below: 5
```