Name: Section:

STAT 113, Spring 99 – Midterm 1

Note: On all problems, please show your work. Just the correct answer without justification and intermediate results is not acceptable.

1. The governor of California proposes a raise for all state employees. Let $\bar{x}$ denote the average monthly salary of state employees (sample mean), and let s_X^2 denote the sample variance.

1a [4pts] If the governor proposes a flat raise of \$70 per month, how would this change the mean and variance? Let $\bar{y}$ and s_Y^2 denote the sample mean and variance after the raise.

Find $\bar{y}$ and s_Y^2 (as functions of $\bar{x}$ and s_X^2).

$$\bar{y} = \frac{1}{n}\sum y_i = \frac{1}{n}\sum(x_i + 70) = \bar{x} + 70$$

and

$$s_Y^2 = \frac{1}{n-1}\sum(y_i - \bar{y})^2 = \frac{1}{n-1}\sum(x_i + 70 - \bar{x} - 70)^2 = s_X^2.$$

1b.* [4pts] What would a 5% increase in the salaries, across the board, do to mean and variance? Let $\bar{z}$ and s_Z^2 denote the sample mean and variance after the raise.
Find $\bar{z}$ and s_Z^2 (as functions of $\bar{x}$ and s_X^2).

$$\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i = \frac{1}{n}\sum(1.05\, x_i) = 1.05\,\frac{1}{n}\sum x_i = 1.05\,\bar{x},$$

and

$$s_Y^2 = \frac{1}{n-1}\sum(y_i - \bar{y})^2 = \frac{1}{n-1}\sum(1.05\, x_i - 1.05\,\bar{x})^2 = 1.05^2 s_X^2.$$

2a. [4pts] Show that $P(E \cap F|G) = P(E|F \cap G)P(F|G)$, for any events E, F and G for which the probabilities involved are defined.

$$P(E \cap F|G) = \frac{P(E \cap F \cap G)}{P(G)}$$

and

$$P(E|F \cap G)P(F|G) = \frac{P(E \cap F \cap G)}{P(F \cap G)} \frac{P(F \cap G)}{P(G)} = \frac{P(E \cap F \cap G)}{P(G)}$$

2b. A card is drawn at random from a deck of playing cards, and another is drawn at random from those that remain. A deck of cards has 52 cards with 13 in each of the four suites (spades, hearts, diamonds, and clubs).
Let A and B denote the events $A =$ "first card is a heart", and $B =$ "2nd card is a heart". Find the probability that

(i) [2pts] both are hearts, i.e. $P(A \cap B)$;

$$P(A \cap B) = P(A)\ P(B|A) = 1/4\ 12/51 = 3/51.$$

(ii) [2pts] the first is not a heart but the second is a heart, i.e. $P(A^c \cap B)$;

$$P(A^c \cap B) = P(A^c)\ P(B|A^c) = 3/4\ 13/51 = 13/17.$$

(iii) [2pts] the second is a heart, i.e., $P(B)$.

$$P(B) = 13/52$$

3. Suppose the number of freshmen, sophomores, juniors, and seniors in a certain college are in the proportion 6:5:4:3, respectively. Women make up 15% of the freshmen, 20% of the sophomores, 30% of the juniors, and 35% of the seniors.

Hint: In your solution please use the following notation:
F = "freshman"; S= "sophomore"; J= "junior"; E= "senior"; W= "woman".

3a. [4pts] What fraction of *all* students are women?

Note $P(F) = 6/18$, $P(S) = 5/18$, $P(J) = 4/18$, $P(E) = 3/18$, $P(W|F) = 0.15$, $P(W|S) = 0.20$, $P(W|J) = 0.30$, and $P(W|E) = 0.35$. Therefore

$$\begin{aligned} P(W) &= P(W|F)\ P(F) + P(W|S)\ P(S) + P(W|J)\ P(J) + P(W|E)\ P(E) \\ &= 1/18(0.15\ 6 + 0.20\ 5 + 0.30\ 4 + 0.35\ 3 = 0.23 \end{aligned}$$

3b. [4pts] What fraction of the women are freshmen?

By Bayes' theorem

$$P(F|W) = \frac{P(W|F)P(F)}{P(W)} = \frac{0.15\ 6/18}{0.23} = 0.22.$$

4. [4pts] The probability that a white male born in 1973 will live to age 40 is .97, and the probability that he will live to age 65 is .66. Find the probability that if he reaches age 40, he will live to be 65.

Hint: In your solution please use the following notation:
A = "lives to age 40"; B = "lives to age 65".

Note that $P(A) = 0.97$, $P(B) = 0.66$, and $A \cap B = B$. Therefore

$$P(B|A) = \frac{P(A \cap B)}{P(A)} = \frac{P(B)}{P(A)} = \frac{0.66}{0.97} = 0.68.$$

5. Assume patients are likely to cancel dentist appointments with probability p, and assume that patient cancelations are independent.

5a. [4pts] In a group of 12 patients scheduled for dental appointments tomorrow morning, let X denote the number of cancellations. Find the probability distribution $p(x)$ of X.

Hint: Since you do not know p you can not evaluate $p(x)$ – an expression for $p(x)$ as a function of x and p is okay.

X is a Binomial r.v. with $n = 12$ and probability of success p.

$$p(x) = (n \;\; x)\, p^x (1-p)^{(n-x)}$$

5b. [4pts] Assume the dentist continues to call patients until she has treated 5 patients, i.e., until she has found the 5th patient who does *not* cancel. Let Y denote the number of patients she has to call (including this 5th treated patient).

Find the probability distribution $p(y)$ of Y.

Hint: Since you do not know p you can not evaluate $p(y)$ – an expression for $p(y)$ as a function of y and p is okay.

Y is a Negative Binomial r.v. with $r = 5$ and probability of success $q = 1 - p$ (note that we are calling until we find the 5-th patient who does *not* cancel).

$$p(y) = (y-1 \;\; r-1)\, q^r (1-q)^{(y-r)} = (y-1 \;\; r-1)\, (1-p)^r p^{(y-r)}$$

Notation: I will use $(N \;\; n)$ to denote the binomial coefficient $\begin{pmatrix} N \\ n \end{pmatrix} = \frac{N!}{(N-n)!\, n!}$

6. An (unknown) number r of the 5 generators in a powerplant have worn rotor shafts. We select at random $n = 3$ of the 5 generators for inspection. Let X denote the number of generators selected for inspection that have worn rotor shafts.

6a. [2pts] What different values can X take on (that is, what is the sample space S)?

$X \in \{x_0, \ldots, x_1\}$ with $x_0 = \max\{0, n - (N - r)\}$ and $x_1 = \min\{r, n\}$.

6b. [4pts] Obtain the probability distribution $p(x)$ of X.

Hint: Since you don't know r you cannot actually evaluate $p(x)$; an expression for $p(x)$ as a function of x and r is ok.

Let $N = 5$ denote the number of generators; and $n = 3$ the sample size. X is a hypergeometric r.v. with

$$p(x) = \frac{(r \;\; x)\,(N - r \;\; n - x)}{(N \;\; n)}$$

6c. [2pts] Calculate the mean and the variance of X.

Hint: Since you don't know r you cannot actually evaluate $E(X)$ and $Var(X)$; the answer is an expression as a function of r.

Using the formulas for the hypergeometric distribution we find
$E(X) = \frac{nr}{N}$; and $Var(X) = \frac{r(N-r)n(N-n)}{N^2(N-1)}$.

6d. We observe $X = 2$ worn rotor shafts among the $n = 3$ inspected generators.

(i) [1pt] What different values can r take on (if $X = 2$)?

Hint: Note that r has to be an integer.

$r \in \{2, 3, 4\}$. Note that $r = 5$ is not possible since we observed $n - X = 1$ rotor shaft which is not rotten, i.e., at most $r = 4$.

(ii) [1pt] Assuming $r = 2$, find the probability $P(X = 2)$ of observing $X = 2$ worn rotors.

$$p(2) = \frac{(2 \;\; 2)\,(3 \;\; 1)}{(5 \;\; 3)} = \frac{1 \cdot 3}{(5 \;\; 3)} = 3/10.$$

(iii)* [2pts] Find the value $\hat{r}$ of r for which $X = 2$ is most likely to be observed.
If the solution is not unique state so and give *all* possible values for $\hat{r}$.

Note that $2 \le r \le 4$ (because we observed $X = 2$ defect rotors and $n - X = 1$ non-defect rotor shaft in the sample). Also

$$\begin{aligned}
\text{For } r = 2: \quad p(2) &= \tfrac{(2\;\;2)\,(3\;\;1)}{(5\;\;3)} = \tfrac{1 \cdot 3}{(5\;\;3)} \\
\text{For } r = 3: \quad p(2) &= \tfrac{(3\;\;2)\,(2\;\;1)}{(5\;\;3)} = \tfrac{3 \cdot 2}{(5\;\;3)} \\
\text{For } r = 4: \quad p(2) &= \tfrac{(4\;\;2)\,(1\;\;1)}{(5\;\;3)} = \tfrac{6 \cdot 1}{(5\;\;3)}
\end{aligned}$$

Since $p(2)$ is highest for both, $r = 3$ and $r = 4$, we can choose $\hat{r} = 3$ or $\hat{r} = 4$.