Name: Section:

STAT 113, Spring 99 – Midterm 2

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

1. Let (X, Y) have the bivariate p.d.f. $f(x, y) = c$ in the portion of the first quadrant bounded by $x + y = 1$, i.e.,

$$f(x,y) = \begin{cases} c & \text{if } x \geq 0,\ y \geq 0 \text{ and } x + y \leq 1, \\ 0 & \text{otherwise} \end{cases}$$

1a. [2pts] Find c.

Hint: A sketch of the set $A = \{x \geq 0,\ y \geq 0 \text{ and } x + y \leq 1\}$ will help (it is simply the first quadrant bounded by $x + y = 1$).

$$1.0 = \int_0^1 \underbrace{\int_0^{1-x} c\, dy}_{=c(1-x)}\ dx = c(x - x^2/2)\big|_0^1 = c/2 \quad \Rightarrow c = 2.0$$

Note: You don't really need to do the integral. Just argue that the p.d.f. is constant and $\{x \geq 0,\ y \geq 0 \text{ and } x + y \leq 1\}$ is a triangle of area $1/2 \Rightarrow \int\int f(x,y) dx\, dy = c\, 1/2$

1b. [2pts] Find the marginal p.d.f. $f(x)$.

$$f(x) = \int_0^{1-x} 2\, dy = 2(1-x), \quad 0 \leq x \leq 1$$

1c. [2pts] Find the conditional p.d.f. $f(y|x)$.

$$f(y|x) = \frac{f(x,y)}{f(x)} = \begin{cases} 1/(1-x) & \text{if } 0 \leq y \lambda 1 - x \\ 0 & \text{otherwise} \end{cases}$$

1d. [2pts] Find $E(X)$.

$$E(x) = \int_0^1 x\, 2(1-x)\ dx = 2\ (x^2/2 - x^3/3)\big|_0^1 = 1 - 2/3 = 1/3.$$

1e. [2pts] Are X and Y independent? Why/why not?

No, X and Y are not independent because

$$f(x,y) = 2 \neq f(x)\ f(y) = 2(1-x)\ 2(1-y).$$

2. A candy bar wrapper says "Net weight 1.4oz." The bars actually vary in weight. To be reasonably confident that most bars weigh at least 1.4 oz., a manufacturer may adjust the production so that the mean weight is 1.5 oz. Assume weights are approximately normally distributed with standard deviation $0.05oz$.

Note: In your solution, denote with X the weight of a candy bar.

2a. [2pts] Find a transformation of X, such that the transformed random variable is standard normal $N(0,1)$.

$Z = (X - \mu)/\sigma$ with $\mu = 1.5$ and $\sigma = 0.05$.

2b. [4pts] Find the proportion of bars weighting less than the advertised 1.4 oz.

The weight X of a candy bar is $N(\mu, \sigma)$ distributed with $\mu = 1.5$ and $\sigma = 0.05$. Therefore

$$P(X \leq 1.4) = P\left(\frac{X - 1.5}{0.05} \leq \frac{1.4 - 1.5}{0.05}\right) = P(Z \leq -2) = 0.5 - 0.4772 = 0.02.$$

2c*. [2pts] Find the weight μ to which the manufacturer has to adjust the production to make sure that less than 1% of the bars weigh less than 1.4 oz.

Find μ such that

$$P(X \leq 1.4) = P\left(\frac{X - \mu}{0.05} \leq \frac{1.4 - \mu}{0.05}\right) = P(Z \leq \underbrace{\frac{1.4 - \mu}{0.05}}_{c}) = 0.01.$$

From Table 4 (page 1094 in the book) we find $c = 2.33 \Rightarrow$

$$c = -2.33 = \frac{1.4 - \mu}{0.05} \quad \Rightarrow \quad \mu = 1.4 - 0.05\,2.33 = \ldots.$$

3. A pharmaceutical company is developing a new drug for stroke patients. Recovery after stroke is measured in Scandinavian Stroke Score (SSS) points. Denote with X_i the SSS points for a patient treated with the new drug and let $\mu = E(X_i)$ and $\sigma^2 = Var(X_i)$ denote expected value and variance. The drug is considered effective if $\mu \geq 4$. In a clinical trial n patients are treated with the new drug. Let $X_1, \ldots, X_n$ denote the measurements for these n patients. Let $\bar{X}$ denote the sample mean.

3a. [2pts] Find $E(\bar{X})$ and $Var(\bar{X})$. (*Note:* Since we do not know μ and σ, the answer will be an expression with μ and σ, not a number.)

$E(\bar{X}) = \mu$ and $Var(\bar{X}) = \sigma^2/n$.

3b. [1pt] Assuming that the X_i are normally distributed, find the sampling distribution $f(\bar{x})$ for $\bar{X}$.

$\bar{X}$ is normally distributed, $\bar{X} \sim N(\mu, \sigma^2/n)$.

3c. [3pts] Assume $\sigma = 10$ and $\mu = 4$. For $n = 100$, find $P(\bar{X} > 1.5)$.

Note that $E(\bar{X}) = \mu = 4$ and $Var(\bar{X}) = \sigma^2/n = 1.0$.

$$P(\bar{X} > 1.5) = P\left(\frac{X-4}{1} > \frac{1.5-4}{1}\right) = P(Z > -2.5) = 0.5 - P(0 \leq Z \leq 2.5) = 0.5 - 0.4938 = 0.9938$$

3d. [2pts] For your answer to **3b** and **3c**, did you need to assume that the X_i be normally distributed? Why/why not?

No. For large n the sampling distribution of $\bar{X}$ is normal even if the probability distribution of X_i is not normal (Central Limit Theorem).

3e*. [2pts] Again assume $\sigma = 10$ and $\mu = 4$.
Find the minimum sample size n such that $P(\bar{X} > 1.5) > 0.95$.

Note that $E(\bar{X}) = \mu = 4$ and $Var(\bar{X}) = \sigma^2/n = 100/n$.

$$P(\bar{X} > 1.5) = P\left(\frac{X-4}{10/\sqrt{n}} > \frac{1.5-4}{10/\sqrt{n}}\right) = P(Z > \underbrace{\frac{1.5-4}{10/\sqrt{n}}}_{c}) = 0.95.$$

From Table 4 (p1094) we find that $c = -1.65 \Rightarrow$

$$\frac{1.5-4}{10/\sqrt{n}} = -1.65 \Rightarrow n \geq (10 \cdot 1.65/2.5)^2 = 43.56$$

i.e., $n \geq 44$.

4. Consider again the SSS measurements X_i, $i = 1, \ldots, n$, for the $n = 100$ stroke patients discussed in problem **3.** Assume $X_i \sim N(\mu, \sigma^2)$. Both parameters, μ and σ, are unknown.

Note: You do not need to have solved problem **3** to work on this problem.

4a. [3pts] Assuming that $\sigma^2 = 100$, find

$$Pr(s^2 \geq 121)$$

Note that $\chi^2 = (n-1)s^2/\sigma^2 \sim \chi^2_{n-1}$. Using Table 8 (p1100) we find:

$$P(s^2 \geq 121) = P\left(\frac{(n-1)s^2}{\sigma^2} \geq \frac{(n-1)121}{\sigma^2}\right) = P(\chi^2 > 119.79) \approx 10\%$$

4b. [2pts] Find the sampling distribution $f(y)$ of $y = \frac{\bar{x}-\mu}{s/\sqrt{n}}$

(state the name of the distribution and values of the relevant parameters).

$t = \bar{x} - \mu/s/\sqrt{n}$ is t-distributed with $\nu = n - 1$ degrees of freedom.

4c. [3pts] For $n = 100$ patients we observe sample mean $\bar{x} = 1.7$ and sample variance $s^2 = 121.0$. Find a cutoff a such that

$$P(\mu \geq a) = 0.90$$

(Remember that σ^2 is *unknown*).

From Table 7 (p1099) we find $P(t \leq 1.29) = 0.90$ (there is no line for $\nu = 99$ – use the cutoff for $\nu = 120$ degrees of freedom). Therefore

$$P\left(\frac{\bar{x}-\mu}{s/\sqrt{n}} \leq 1.29\right) = 0.9$$

and

$$P(\mu \geq \bar{x} - 1.29s/\sqrt{n}) = P(\mu \geq 1.7 - 1.29 \cdot 1.1) = 0.9$$

4d*. [2pts] For $n = 100$ patients we observe sample mean $\bar{x} = 1.7$ and sample variance $s^2 = 121.0$. Find cutoffs c and d such that

$$Pr(c \leq \sigma^2 \leq d) = 0.95$$

and $Pr(\sigma^2 \geq c) = 0.025$.

For $\chi^2 \sim \chi^2_{99}$ we find from Table 8 (p1100)

$$P(74.22 \leq \chi^2 \leq 129.6) \approx 0.95 \Rightarrow \ldots \Rightarrow P((n-1)s^2/129.6 \leq \sigma^2 \leq (n-1)s^2/74.22) \approx 0.95$$

i.e., $c = 99 \cdot 121/129.6$ and $d = 99 \cdot 121/74.22$.