

STAT 324 Sample midterm problems

1. $1/7$ of men and $1/9$ of women are color-blind. 60% of population are females.

(a) A person is chosen at random and that person is color-blind. What is the probability that the person is male?

Solution. Let C be the event the person is color blind. Let M and F be the events the person is male and female respectively. Then we need to compute

$$P(M|C) = \frac{P(M \cap C)}{P(C)} = \frac{P(C|M)P(M)}{P(C|M)P(M) + P(C|F)P(F)} = \frac{\frac{1}{7} \cdot \frac{4}{10}}{\frac{1}{7} \cdot \frac{4}{10} + \frac{1}{9} \cdot \frac{6}{10}}.$$

(b) A person is chosen at random and that person is male. What is the probability that the person is color-blind?

Solution. $P(C|M) = \frac{1}{7} \cdot \frac{4}{10}$.

2. The life time X of an electronic component follows the distribution $f(x) = e^{-x}$ for $x > 0$ and $f(x) = 0$ otherwise. Three of these components operate independently in a piece of equipment. The equipment fails if at least two of the components fail. Find the probability that the equipment will operate for at least 2 hours without failure.

Solution. The probability of a single component failure is

$$p = P(X < 2) = \int_0^2 e^{-x} dx = 1 - e^{-2}.$$

The equipment will fail if at least two out of 3 components fail. So the probability of the equipment failure is

$$\binom{3}{2}p^2(1-p) + \binom{3}{3}p^3 = 3p^2(1-p) + p^3.$$

The probability of the equipment operating for at least 2 hours is then

$$1 - 3p^2(1-p) - p^3.$$

3. The achievement scores for a college entrance examination are normally distributed with mean 70 and variance 10. What fraction of the scores lies between 75 and 80? You may use the following R output.

```
> pnorm(c(1:4), 1, 2)
[1] 0.50 0.69 0.84 0.93
```

Solution. Let X be the achievement score and let $Z = \frac{X-70}{\sqrt{10}} \sim N(0, 1)$. Need to find

$$p = P\left(\frac{75-70}{\sqrt{10}} < Z < \frac{80-70}{\sqrt{10}}\right) = P(0.5 < Z < 1).$$

However, the R output is given in terms of $W \sim N(1, 2^2)$. The relationship between W and Z is $W = 2Z + 1$. So

$$\begin{aligned} p &= P(2 \cdot 0.5 + 1 < W < 2 \cdot 1 + 1) = P(2 < W < 3) \\ &= \text{pnorm}(3, 1, 2) - \text{pnorm}(2, 1, 2) = 0.84 - 0.69 = 0.15. \end{aligned}$$

4. The density function of the random variable X is $f(x) = \frac{x}{2}$ for $0 < x < 2$ and $f(x) = 0$ otherwise.
 (a) Find the cumulative distribution function of X and plot its graph.

Solution. $F_X(x) = \int_0^x \frac{x}{2} dx = \frac{x^2}{4}$ for $0 < x < 2$, $F_X(x) = 0$ for $x \leq 0$ while $F_X(x) = 1$ for $x \geq 2$.

- (b) Derive the probability density function of $Y = 2X + 1$ using the cumulative distribution function of X .

Solution. $F_Y(y) = P(Y \leq y) = P(2X + 1 \leq y) = F_X(\frac{y-1}{2}) = \frac{1}{4} \left(\frac{y-1}{2} \right)^2$ for $1 < y < 5$. $f_Y(y) = \frac{dF_Y(y)}{dy} = \frac{1}{6} \frac{y-1}{2}$ for $1 < y < 5$.

5. Let X and Y have joint probability density

$$f(x, y) = cx, \text{ for } 0 < x < 1, 0 < y < 1 \text{ and } 1 < x + y < 2$$

and $f(x, y) = 0$ otherwise.

- (a) Find constant c .

Solution. Draw the domain where the joint density is defined.

$$1 = \int_0^1 \int_{1-y}^1 cx \, dx \, dy = \frac{c}{2} \int_0^1 (1 - (1-y)^2) \, dy = \frac{c}{2} \left(y + \frac{1}{3}(1-y)^3 \right) \Big|_0^1 = \frac{c}{2} \left(1 - \frac{1}{3} \right).$$

So $c = 3$.

- (b) Find the marginal distributions of X and Y .

Solution. $f_X(x) = \int_{1-x}^1 3x \, dy = 3x^2$ for $0 < x < 1$. $f_Y(y) = \int_{1-y}^1 3x \, dx = \frac{3}{2}(1 - (1-y)^2)$ for $0 < y < 1$.

- (c) Find the conditional probability $P(X < 1 | X + Y > 1)$.

Solution. This is trivially

$$\frac{P(X < 1, X + Y > 1)}{P(X + Y > 1)} = 1$$

6. Let $X_1, \dots, X_n$ be a random sample from Bernoulli distribution with parameter p . (a) What is $\mathbb{E}(S^2/p^2)$? S^2 is the sample variance. Explain your results.

Solution. The sample variance is an unbiased estimator of the population variance. Hence $\mathbb{E}(S^2/p^2) = \mathbb{E}(S^2)/p^2 = \mathbf{Var}(X_i)/p^2$. The variance for a Bernoulli random variable can be easily computed as $\mathbf{Var}(X_i) = p(1-p)$. So $\mathbb{E}(S^2/p^2) = (1-p)/p$.

- (b) Find an unbiased estimator for p^2 . Explain your results.

Solution. There are many solutions to this problem. We know S^2 and $\bar{X}$ will be unbiased estimators of population variance $p(1-p)$ and mean p respectively. So $\mathbf{E}(\bar{X}) - \mathbf{E}(S^2) = p - p(1-p) = p^2$. Hence, $\bar{X} - S^2$ is an unbiased estimator for p^2 .

7. Let X_1, X_2 be a random sample from $N(0, 1/\theta)$. Note that the sample size is 2 and the density function for X_i is $f(x_i) = \frac{\sqrt{\theta}}{\sqrt{2\pi}} \exp(-\theta x_i^2/2)$. Find the likelihood function and use it to obtain the maximum likelihood estimator of θ .

Solution. The likelihood function is $L(\theta) = \theta/(2\pi) \exp(-\theta(x_1^2 + x_2^2)/2)$. Get log-likelihood function $\ln L(\theta) = \text{const} + \ln \theta - \theta(x_1^2 + x_2^2)/2$. Differentiate with respect to θ to get the maximum:

$$\frac{d \ln L(\theta)}{d\theta} = \frac{1}{\theta} - \frac{1}{2}(x_1^2 + x_2^2) = 0.$$

Solving the equation, we get $\hat{\theta} = 2/(x_1^2 + x_2^2)$.