2028: Basic Statistical Methods
Solutions - Homework 2

1 a. $P(\text{card is heart 'A'}|\text{the card is red}) = \frac{P(\text{card is heart 'A' \& the card is red})}{P(\text{the card is red})} = \frac{\frac{1}{52}}{\frac{26}{52}} = \frac{1}{26}$

b. $P(\text{card is heart}|\text{the card is red}) = \frac{P(\text{card is heart \& the card is red})}{P(\text{the card is red})} = \frac{\frac{13}{52}}{\frac{26}{52}} = \frac{1}{2}$

c. $P(\text{card is King}|\text{the card is red}) = \frac{P(\text{card is King\& the card is red})}{P(\text{the card is red})} = \frac{\frac{2}{52}}{\frac{26}{52}} = \frac{1}{13}$

d. $P(\text{card is King}) = \frac{4}{52} = \frac{1}{13}$
$P(\text{card is King}) = P(\text{card is King}|\text{the card is red})$. In other words, these two events are statistically independent.

2 Let X be the time that elapses between the end of the hour and the end of the lecture.

(a) f is the pdf of X thus the area under the graph of f is 1:

$$1 = \int_{-\infty}^{\infty} f(x)\,dx = \int_0^2 kx^2 = k\left(\frac{x^3}{3}|_2 - \frac{x^3}{3}|_0\right) = k\frac{8}{3}.$$

Thus $k = \frac{3}{8}$.

(b) We need $P(X \leq 1)$:

$$P(X \leq 1) = P(0 \leq X \leq 1) = \int_0^1 \frac{3}{8}x^2 = \frac{1}{8}\left(x^3|_1 - x^3|_0\right) = \frac{1}{8} = .125.$$

(c)

$$P(1 \leq X \leq 1.5) = \int_1^{1.5} \frac{3}{8}x^2 = \frac{1}{8}\left(x^3|_{1.5} - x^3|_1\right) = \frac{1}{8}\left(\frac{3}{2}\right)^3 - \frac{1}{8}(1)^3 = \frac{19}{64} = .2969$$

(d)

$$P(X \geq 1.5) = 1 - \int_0^{1.5} \frac{3}{8}x^2 = 1 - \frac{1}{8}\left(x^3|_{1.5} - x^3|_0\right) = 1 - \frac{1}{8}\left(\frac{3}{2}\right)^3 = \frac{37}{64} = .5781$$

(e) The cdf of X is for $x \geq 0$:

$$F(x) = \int_{-\infty}^{x} f(t)dt = \int_0^x \frac{3}{8}t^2 dt = \frac{1}{8}\left(t^3|_a - t^3|_0\right) = \frac{x^3}{8}$$

Thus the cdf of X is:

$$F(x) = \begin{cases} \frac{x^3}{8} & \text{if } 0 \leq x \leq 2 \\ 0 & \text{otherwise} \end{cases}$$

The probability in (b) is:

$$F(1) - F(0) = \frac{1^3}{8} - \frac{0^3}{8} = \frac{1}{8} = .125.$$

The probability in (c) is:

$$F(1.5) - F(1) = \frac{1.5^3}{8} - \frac{1^3}{8} = \frac{3.375 - 1}{8} = .2969.$$

The probability in (d) is:

$$F(2) - F(1.5) = \frac{2^3}{8} - \frac{1.5^3}{8} = 1 - \frac{3.375}{8} = .5781.$$

3 The answers are as follows:
(a) The binomial assumption seems reasonable. First, each animal may have or do not have the trait and therefore the outcome is dichotomous. Second, it's reasonable to assume that the probability of an animal having the trait is not affected by whether or not the other animals have the trait, if they are not come from the same family. Therefore the independence assumption seems reasonable. Third, the problem assumes that the animals have the same probability(0.2) of having the trait. Therefore the homogeneity assumption holds.
(b) The binomial assumption seems not reasonable. First, the experiment was performed using a single rat. The rat would probably learn how to get the food in the process and so the probability of getting the food at later trials would be higher. Therefore the homogeneity assumption seems to be violated. Second, because given that the rat was successful in getting the food at a particular trial(say the 9th trial), the probability that the rat would be succeeded on the 10th trial would be higher than if we didn't know the outcome of the previous (the 9th) trial. Therefore the independence assumption seems to be violated.

4 Let X be the number of neurons that have higher firing rates for Condition 2. Then we assume $X \sim B(25, .38)$.

(a) We want $P(X = 7)$. This is the value of the binomial pdf $p(7)$. The R command used is

```
dbinom(7,25,.38)
```

We get $P(X = 7) = 0.1007$.

(b) We want $P(X \geq 8)$. This is $1 - P(X \leq 7) = 0.7932$. The R command used is

```
1 - pbinom(7,25,.38)
```

(c) We want $P(X \leq 8)$. This is the value of the binomial cdf $F(8)$. The R command used is

```
pbinom(8,25,.38)
```

We get $P(X \leq 8) = 0.3458$.

(d) We want $P(X > 8)$. We use $P(X > 8) = 1 - P(X \leq 8)$, and $P(X \leq 8) = 0.3458$ from part (c). Thus $P(X > 8) = 0.6542$. The R command used is

```
1 - pbinom(8,25,.38)
```

(e) We want $P(8 \leq X \leq 10) = P(X \leq 10) - P(X \leq 7) = 0.4577$ The R command used is

```
pbinom(10,25,.38)-pbinom(7,25,.38)
```

5 Let X denote the number of stars in 16 cubic light-years, then $X \sim Poisson(\lambda, T), where \lambda = 1/16, T = 16, \lambda T = 1$.

a. $P(X \geq 2) = 1 - P(X=0) - P(X=1) = 1 - \frac{e^{-1}1^0}{0!} - \frac{e^{-1}1^1}{1!} = 1 - \frac{2}{e}$

b. Let Y denote the number of stars in T_1 cubic light-years, then $Y \sim Poisson(\lambda, T_1), where \lambda = 1/16, \lambda T = \frac{T_1}{16}$.

$P(Y \geq 1) > 0. \iff 1 - P(Y=0) > 0.95 \iff P(Y=0) < 0.05 \iff e^{-\frac{T_1}{16}} < 0.05$ $\iff T_1 > 16log(20)$cubic light-years

6 a. Let X denote the diameter of a drilled hole in mm, then $X - 5 \sim Exp(10)$.
$E(X) = E(X-5) + 5 = 5.1, Var(X) = Var(X-5) = 0.01$

b. $P(X > 5.1) = P(X - 5 > 0.1) = 1 - P(X - 5 \leq 0.1) = 1 - (1 - \frac{1}{e}) = \frac{1}{e}$

7 a. Let X be the number of aircraft arriving within an hour. Then X $\sim$ Piosson(1).
$P(X > 3) = 1 - P(X=0) - P(X=1) - P(X=2) - P(X=3) = 1 - \frac{8}{3e} = 0.019$

b. $P(X \leq 3) = \frac{8}{3e}$ $P($ no interval contains more than three arrivals$) = P($ all 30 intervals contain $\leq 3) = P(X \leq 3)^{30} = 0.56$

c. Let Y be the time between arrivals of small aircraft. Then Y $\sim$ Exp(1).
$P(Y > y) = 0.1 \Leftrightarrow e^{-y} = 0.1 \Leftrightarrow y = log10 = 2.30$

8 a. Because X is normally distributed then Y is also normally distributed with mean and variance derived as follows:

$$\mathbb{E}(Y) = 0.8\mathbb{E}(X) + 20 = 0.8 * 75 + 20 = 80$$
$$\mathbb{V}(Y) = 0.8^2\mathbb{V}(X) = 0.8^2 * 25 = 16$$

b. $P(Y > 75) = 1 - P(Y \leq 75) = 1 - 0.105 = 0.895$

c. Need to find $q_{.25}$ such that $P(Y \leq q_{.25}) = .25$. After standardization and using the normal tables, we find that $q_{.25} = 77.302$

9 a. Let X denote the yearly income of households. Since 400 is large, we could see this as a normal distribution according to the central limit theorem. Then $\bar{X} \sim N(25000, \frac{(10000)^2}{400}) = N(25000, (500)^2)$
$P(24000 \leq X < 25500) = P(\frac{24000-25000}{500} \leq \frac{X-25000}{500} < \frac{25500-25000}{500}) = P(-2 \leq Z < 1) = 0.82$

b. We are not given any information about the distribution of X.

10 Let X be the melting temperature in degree Celcius and Y be the melting temperature in degree Fahrenheit. Then $\mu_X = 150, \sigma_X = 2$ and $Y = 1.8X + 32$.

$$\begin{aligned} \mu_Y &= \mu_{1.8X+32} = 1.8\mu_X + 32 = 1.8 \times 150 + 32 = 302 \\ \sigma_Y &= \sigma_{1.8X+32} = |1.8|\sigma_X = 1.8 \times 2 = 3.6 \end{aligned}$$