# Assignment 1

## Exercise 1

We have the following:

$P(A) = \frac{1}{2}$ , $P(B'|A) = \frac{3}{4}$ and $P(A \cup B) = \frac{3}{4}$

a) $P(A \cap B) = ;$

$P(B'|A) = 1 - P(B|A) = \frac{3}{4} \Rightarrow \boxed{P(B|A) = \frac{1}{4}}$

$P(A \cap B) = P(B|A) \cdot P(A) = \frac{1}{4} \cdot \frac{1}{2} = \frac{1}{8} \Rightarrow \boxed{P(A \cap B) = \frac{1}{8}}$

b) $P(B) = ;$

$P(A \cup B) = P(A) + P(B) - P(A \cap B) = \frac{1}{2} + P(B) - \frac{1}{8} \quad (1)$

$(1) \Rightarrow P(A \cup B) = \frac{3}{4} \Leftrightarrow$

$\frac{1}{2} + P(B) - \frac{1}{8} = \frac{3}{4} \Leftrightarrow$

$P(B) = \frac{3}{4} - \frac{1}{2} + \frac{1}{8} \Leftrightarrow$

$P(B) = \frac{6 - 4 + 1}{8} \Leftrightarrow$

$\boxed{P(B) = \frac{3}{8}}$

c) $P(A|B) = ;$

$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{\frac{1}{8}}{\frac{3}{8}} = \frac{1}{3} \Rightarrow \boxed{P(A|B) = \frac{1}{3}}$

d) Two events A and B are independent if and only if
$P(A \cap B) = P(A) \cdot P(B)$

$$P(A \cap B) = \frac{1}{8} \quad (1)$$

$$\left.\begin{array}{l} P(A) = \frac{1}{2} \\ P(B) = \frac{3}{8} \end{array}\right\} \Rightarrow P(A) \cdot P(B) = \frac{3}{16} \quad (2)$$

$$\Rightarrow (1) \neq (2)$$

So, the events A and B are not independent.

## Exercise 2

We have a fair six-sided die and we roll it until we obtain a 6. X denotes the number of rolls/tries required to obtain a 6.

a) The random variable X follows the geometric distribution which takes values in the set $\{1, 2, \ldots, n\}$. The probability of success $p$ is $\frac{1}{6}$.

The mass function $P$ is: $P(X=n) = p \cdot (1-p)^{n-1}$ $(=)$

$$\boxed{P(X=n) = \frac{1}{6} \cdot \left(\frac{5}{6}\right)^{n-1}}$$

b) For $X = 10 \Rightarrow$

$$P(X=10) = \frac{1}{6} \cdot \left(\frac{5}{6}\right)^9 = \frac{5^9}{6^{10}} = 0{,}0323 \Rightarrow$$

$$\Rightarrow \boxed{P(X=10) = 0{,}0323}$$

c)

The expected value comes from the formula $\frac{1}{p}$

$$E[X] = \frac{1}{p} = \frac{1}{\frac{1}{6}} = 6 \Rightarrow \boxed{E[X] = 6}$$

## Exercise 3

We have a Communication channel which sents a symbol "1" with probability $P(\text{sent } 1) = 0.4$ and a Symbol "0" with probability $P(\text{sent } 0) = 0.6$

We have the following conditional probabilities:

$P(\text{recieve } 1 | \text{sent } 0) = 0.1$

$P(\text{recieve } 0 | \text{sent } 1) = 0.2.$

a) The scope is to find the value of $P(\text{Sent } 1 | \text{receive } 1)$

$$P(\text{Sent } 1 | \text{recieve } 1) = p(\text{recieve } 1 | \text{sent } 1) \cdot \frac{P(\text{sent } 1)}{P(\text{recieve } 1)} \quad (1)$$

$\boxed{P(\text{sent } 1) = 0.4}$ (2)

$P(\text{recieve } 1) = P(\text{sent } 1) \cdot P(\text{receive } 1 | \text{sent } 1) + P(\text{sent } 0) \cdot P(\text{recieve } 1 | \text{sent } 0)$ $\Leftrightarrow$

$P(\text{recieve } 1) = 0.4 \cdot (1 - 0.2) + 0.6 \cdot (0.1)$ $\Leftrightarrow$

$\boxed{P(\text{recieve } 1) = 0.38}$ (3)

$P(\text{recieve } 1 | \text{sent } 1) = 1 - P(\text{receive } 0 | \text{sent } 1) = 1 - 0.2 = 0.8$ (4)

From (2), (3), (4) $\overset{(1)}{\Rightarrow}$ $P(\text{Sent } 1 | \text{receive } 1) = P(\text{receive } 1 | \text{sent } 1) \cdot \frac{P(\text{sent } 1)}{P(\text{receive } 1)}$ $\Leftrightarrow$

$$P(\text{sent } 1 | \text{receive } 1) = \frac{(0.8) \cdot (0.4)}{0.38} \Leftrightarrow$$

$$\boxed{P(\text{Sent } 1 | \text{receive } 1) = 0.8421}$$

b) The scope is to find the probability if two "0" Symbols received the two "0" symbols sents. This is equal to the probability $(P(\text{sent } 0 | \text{receive } 0))^2$.

$$P(\text{sent } 0 \mid \text{receive } 0) = p(\text{receive } 0 \mid \text{sent } 0) \cdot \frac{P(\text{sent } 0)}{P(\text{receive } 0)} \quad (1)$$

$$P(\text{sent } 0) = 0.6 \quad (2)$$

$$P(\text{receive } 0 \mid \text{sent } 0) = 1 - P(\text{receive } 1 \mid \text{sent } 0) = 1 - 0.1 = 0.9. \quad (3)$$

$$P(\text{receive } 0) = p(\text{sent } 0) \cdot p(\text{receive } 0 \mid \text{sent } 0) + p(\text{sent } 1) \cdot p(\text{receive } 0 \mid \text{sent } 1) \Leftrightarrow$$

$$P(\text{receive } 0) = (0.6) \cdot (0.9) + (0.4) \cdot (0.2) \Leftrightarrow$$

$$\boxed{P(\text{receive } 0) = 0.62} \quad (4)$$

From (2), (3), (4) $\overset{(1)}{\Rightarrow}$ $P(\text{sent } 0 \mid \text{receive } 0) = P(\text{receive } 0 \mid \text{sent } 0) \cdot \frac{P(\text{sent } 0)}{P(\text{receive } 0)} \Leftrightarrow$

$$P(\text{sent } 0 \mid \text{receive } 0) = \frac{(0.9) \cdot (0.6)}{0.62} \Leftrightarrow$$

$$\boxed{P(\text{sent } 0 \mid \text{receive } 0) = 0.8709}$$

So, the final probability $P$ is $P = (0.8709)^2 = 0.7584$

# Exercise 4

We have a continuous random variable $X$ with a pdf $f(x)$ and a cdf $F(x)$.
Also, for a fixed value $x_0 \Rightarrow F(x_0) < 1 \Leftrightarrow \int_{-\infty}^{x_0} f(x) < 1$.

$$g(x) = \begin{cases} \dfrac{f(x)}{1-F(x_0)}, & x \geq x_0 \\ 0, & x < 0. \end{cases}$$

Since $f(x)$ is a pdf of $X$ then:

$$f(x) \geq 0, \ \forall x \quad \text{and} \quad \int_{-\infty}^{+\infty} f(x)dx = 1.$$

The scope is to show that $g(x)$ is a valid pdf. The $g(x)$ is a valid pdf if only if:

$$\text{i) } g(x) \geq 0, \ \forall x \quad \text{and} \quad \text{ii) } \int_{-\infty}^{+\infty} g(x)dx = 1.$$

i) For $x < x_0 \Rightarrow g(x) = 0 \geq 0$
for $x \geq x_0 \Rightarrow g(x) = \dfrac{f(x)}{1-F(x_0)} \geq 0$ $\Bigg\} \Rightarrow$ So, $g(x) \geq 0, \forall x$

$\left.\begin{array}{l} f(x) \geq 0, \ \forall x \\ 1-F(x_0) > 0 \end{array}\right\}$

ii)

For $x < x_0 \Rightarrow \int_{-\infty}^{x_0} g(x)dx = \int_{-\infty}^{x_0} 0 dx = 0 \quad (1)$

For $x \geq x_0 \Rightarrow \int_{x_0}^{+\infty} g(x)dx = \int_{x_0}^{+\infty} \dfrac{f(x)}{1-F(x_0)} dx = \dfrac{1}{1-F(x_0)} \int_{x_0}^{+\infty} f(x)dx =$

$= \dfrac{1}{1-F(x_0)} \left(1 - \int_{-\infty}^{x_0} f(x)dx\right) = \dfrac{\cancel{1-F(x_0)}}{\cancel{1-F(x_0)}} = 1. \quad (2)$

From (1) and (2) $\Rightarrow \int_{-\infty}^{+\infty} g(x)dx = \int_{-\infty}^{x_0} g(x)dx + \int_{x_0}^{+\infty} g(x)dx =$

$= 0 + 1 = 1 \Rightarrow \boxed{\int_{-\infty}^{+\infty} g(x)\,dx = 1}$

So,

$g(x) \geq 0, \forall x$ and $\int_{-\infty}^{+\infty} g(x)dx = 1$

That means that $g(x)$ is a valid pdf.

Exercise 5

We have a Poisson random variable $X$ with $E[X] = \lambda = 3 = Var[X]$. It takes values $x \in \{0, 1, 2, \dots\}$

$$f(x|\lambda) = P(X=x|\lambda) = \frac{e^{-\lambda} \cdot \lambda^x}{x!}$$

For $\lambda = 3 \Rightarrow f(x|3) = \frac{e^{-3} \cdot 3^x}{x!}$

a)

For $X=0 \Rightarrow f(X=0|\lambda) = P(X=0|\lambda) = P(X=0|3) = e^{-3} = \frac{1}{e^3} \approx 0.0497$

b)

For $X \geq 3 \Rightarrow f(X \geq 3|3) = P(X \geq 3|3) = 1 - (P(X=0|3) + P(X=1|3) + P(X=2|3)) =$

$= 1 - P(X \leq 2|3) = 1 - 0.4231901 = 0.5768099$

(using R)

c) For length 3 minutes the $\lambda$ becomes 9.

So, $f(x|9) = P(X=x|9) = \frac{e^{-9} 9^x}{x!}$

For $X \leq 5 \Rightarrow f(X \leq 5|9) = P(X \leq 5|9) = 0.1156905$

(using R)

Exercise 6

We have a random sample $X_1, X_2, \ldots, X_n$ from the Weibull$(k, \lambda)$ distribution with $k$ known and $\lambda$ unknown.

The pdf is: $p(x; k, \lambda) = \frac{k}{\lambda}\left(\frac{x}{\lambda}\right)^{k-1} \cdot e^{-\left(\frac{x}{\lambda}\right)^k}$, $x \geq 0, \lambda > 0$

Let's take the joint likelihood function:

$$L(\lambda \mid x) = f(x \mid \lambda) = \prod_{i=1}^{n} f(x_i \mid \lambda)$$

$$\prod_{i=1}^{n} f(x_i \mid \lambda) = \prod_{i=1}^{n} \frac{k}{\lambda}\left(\frac{x_i}{\lambda}\right)^{k-1} \cdot e^{-\left(\frac{x_i}{\lambda}\right)^k} = \left(\frac{k}{\lambda}\right)^n \prod_{i=1}^{n} \frac{x_i^{k-1}}{\lambda^{k-1}} e^{-\left(\frac{x_i}{\lambda}\right)^k} =$$

$$= \left(\frac{k}{\lambda}\right)^n \cdot \left(\prod_{i=1}^{n} x_i^{k-1}\right) \cdot \left(\lambda^{-n(k-1)}\right) \prod_{i=1}^{n} e^{-\left(\frac{x_i}{\lambda}\right)^k} =$$

$$= \frac{k^n}{\lambda^n} \left(\prod_{i=1}^{n} x_i^{k-1}\right) \cdot \left(\lambda^{-n(k-1)}\right) \cdot \left(e^{-\sum_{i=1}^{n}\left(\frac{x_i}{\lambda}\right)^k}\right) =$$

$$= k^n \cdot \prod_{i=1}^{n} x_i^{k-1} \cdot \lambda^{-\cancel{n} - nk + \cancel{n}} \cdot e^{-\sum_{i=1}^{n}\left(\frac{x_i}{\lambda}\right)^k} =$$

$$= k^n \prod_{i=1}^{n} x_i^{k-1} \cdot \lambda^{-n} \cdot e^{-\sum_{i=1}^{n}\left(\frac{x_i}{\lambda}\right)^k} \quad (1)$$

From (1) $\Rightarrow$ $h(x) = k^n \cdot \prod_{i=1}^{n} x_i^{k-1}$ and $g(T(x), \lambda) = \lambda^{-n} \cdot e^{-\sum_{i=1}^{n}\left(\frac{x_i}{\lambda}\right)^k}$

with $T(x) = \sum_{i=1}^{n} x_i^k$. So, $f(x \mid \theta) = h(x) g(T(x), \lambda)$

Based on the factorization theorem if $f(x) = h(x) g(T(x), \lambda)$, $\forall x \geq 0$ and $\lambda > 0$ then the $T(x) = \sum_{i=1}^{n} x_i^k$ is a sufficient statistic of $\lambda$.

## Exercise 1

We have two random samples X1 and X2 with the following:

For sample X1: $\bar{X_1} = 45.3$, $S_1^2 = 4.1$ and $n_1 = 15$

For sample X2: $\bar{X_2} = 41.8$, $S_2^2 = 3.9$ and $n_2 = 18$.

a) Construct a 95% confidence interval for the mean of the population for both samples.

Because, in both samples we do not know the variance of the population we will have the following:

- Sample $X_1$

$$\bar{X_1} - t_{n_1-1, 0.975} \frac{S_1}{\sqrt{n_1}} \leq \mu_1 \leq \bar{X_1} + t_{n_1-1, 0.975} \frac{S_1}{\sqrt{n_1}} \quad (1)$$

$\bar{X_1} = 45.3$

$S_1 = \sqrt{S_1^2} = \sqrt{4.1} \approx 2.0248$

$\sqrt{n_1} = \sqrt{15} \approx 3.8729$

$t_{n_1-1, 0.975} = t_{14, 0.975} = 2.144787$ (Using R)

So, (1) =>
- $45.3 - (2.144787) \cdot \left(\frac{2.0248}{3.8729}\right) \approx 44.1610$
- $45.3 + (2.144787) \cdot \left(\frac{2.0248}{3.8729}\right) \approx 46.4213.$

So, (1) => $\boxed{44.1610 \leq \mu_1 \leq 46.4213}$

- Sample $X_2$:

$$\bar{X}_2 - t_{n_2-1,0.975}\frac{S_2}{\sqrt{n_2}} \le \mu_2 \le \bar{X}_2 + t_{n_2-1,0.975}\frac{S_2}{\sqrt{n_2}} \quad (2)$$

$\bar{X}_2 = 47.8$
$S_2 = \sqrt{S_2^2} = \sqrt{3.9} \approx 1.9748$
$\sqrt{n_2} = \sqrt{18} \approx 4.2426$
$t_{n_2-1,0.975} = t_{17,0.975} = 2.109816$ (using R)

So, $(2) \Rightarrow \cdot\ 47.8 - (2.109816)\cdot\left(\frac{1.9748}{4.2426}\right) \approx 46.8179$

$\cdot\ 47.8 + (2.109816)\cdot\left(\frac{1.9748}{4.2426}\right) \approx 48.7820$

So, $(2) \Rightarrow \boxed{46.8179 \le \mu_2 \le 48.7820}$

b) The significance level is $\alpha = 0{,}05$

The hypothesis test is:

$H_0: \mu_1 = \mu_2$
$H_1: \mu_1 \neq \mu_2$.

Because the samples are independent and variances of their population are unknown and not equal we have:

$$T = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} = \frac{-2.5}{\sqrt{\frac{4.1}{15} + \frac{3.9}{18}}} = \frac{-2.5}{0.7} \approx -3.5714$$

$$v = \frac{\left(S_1^2/n_1 + S_2^2/n_2\right)^2}{\left(\frac{S_1^2}{n_1}\right)^2/(n_1-1) + \left(\frac{S_2^2}{n_2}\right)^2/(n_2-1)} =$$

$$= \frac{\left(\frac{4.1}{15} + \frac{3.9}{18}\right)^2}{\frac{\left(\frac{4.1}{15}\right)^2}{14} + \frac{\left(\frac{3.9}{18}\right)^2}{17}} = \frac{(0.49)^2}{0.0053 + 0.027} = \frac{0.2401}{0.008} \approx 30$$

$$t_{v,\alpha/2} = t_{30,\ 0.025} \approx -2.0422$$
$$-t_{v,\alpha/2} = -t_{30,\ 0.025} \approx 2.0422$$

$T = -3.5714 < t_{30,\ 0.025}$
$T = -3.5714 < -t_{30,\ 0.025}$ $\Rightarrow$ So, we reject the $H_0$ and we can say the means are statistically different.

c) The significance level is 0.01

The hypothesis test is:

$$H_0: \mu_1 = 45$$
$$H_1: \mu_1 > 45$$

Because the variance of the population is unknown we have that:

$$T = \frac{\bar{X}_1 - 45}{\frac{S_1}{\sqrt{n_1}}} = \frac{45.3 - 45}{0.5228} = \frac{0.3}{0.5228} \approx 0.5738$$

$t_{n-1, 1-\alpha} = t_{14, 0.99} \approx 2.6244$

Since, $T < t_{14, 0.99}$ we fail to reject $H_0$, so the mean is not greater than 45 using out the 1% significance level.