

## R Programming assignment:

### Exercise 1

- a) Construct 2 vectors  $x = (i), i = 1, \dots, 200; y = (j), j = 1, \dots, 200$ , then plot the points (x,y)

You may use the R command `Plot(x, y)`.

- b) Use the following command to graph 20 points  
`plot(x,y,xlim=range(0:5),ylim=range(0:5),type='b',main="X vs Y")`

Note: go to help plot to see more about the “plot” arguments

### Exercise 2:

- a) Sketch the probability mass function of  $X \sim \text{binom}(\text{size} = 35, \text{prob} = 0.461)$ .  
b) Sketch the cumulative distribution function  
c) Find the probability that  $X$  is equal to 21.  
d) Find the probability that  $X$  is at most 18.  
e) Find the probability that  $X$  is more than 30  
f) Find the probability that  $X$  is between 12 and 24, inclusive.

**Solve the following using 3 methods as indicated and compare**

- g) Find the mean of  $X: E(X)$ , using the formula  $\mu = \sum xf(x)$   
h) Install and use the package “`distrEx`” to verify your answer in question g), with the command `E(X)` and then by simply computing  $np$   
i) Find the variance and standard deviation of  $X: \sigma^2$  and  $\sigma$ , using the formula  
$$\sigma^2 = E(X^2) - (E(X))^2 = E(X^2) - \mu^2$$
 (or equivalent)  
j) Use the package to verify your answer, using the commands `var(X)` and `sd(X)` and then by simply computing  $npq = np(1-p)$ .

### Exercise 3: experiments with the binomial distribution

Let  $X \sim \text{binom}(n = 100,000, p = 0.01)$ ,  $p = 0.01$  probability of side effect resulting from a treatment and  $n = 100,000$  treated individuals

- a) Use your calculator to find  $100000C110$   
b) Use R to find  $100000C65$   
c) Use R to find  $100000C125$ , then  $100000C1100$   
d) Use `dbinom` to find  $f(1100)$   
e) Use R to compute  $F(980)$   
f) Use the sum operator and `dbinom` to find  $P(900 \leq X \leq 1200)$   
g) Compute  $F(1200) - F(900)$  using `pbinom`.  
h) Use `data.frame()` to construct a table of  $x, F(x)$  for  $x = 980$  to  $x = 1020$   
i) Use `data.frame()` to construct a table of  $x, f(x)$  for  $x = 980$  to  $x = 1020$   
j) **Graphing pmf  $f(x)$**   
Use the following code to graph the indicated binomial distribution for  $p = 0.1, 0.3, 0.5, 0.7$  and  $0.9$ .

```
n <- 200
k <- seq(0, n, by = 1)
```

`plot(k, dbinom(k, n, p))`

**k) Graphing cdf F(x):**

use same as question j) with the command `pbinom`

- l) Generate a vector of 100 number from a binomial distribution  $X \sim \text{binom}(200, 0.6)$   
Use `mean(vector)` to find the average. Compare to the formula of the mean of a binomial distribution. What is the error?

**Exercise 3**

Let  $X \sim \text{Pois}(\lambda = 5)$ :

- i) Use `data.frame(...)` to construct a table of  $x, F(x)$  for  $x = 0$  to  $x = 20$   
ii) Use `data.frame()` to construct a table of  $x, f(x)$  for  $x = 0$  to  $x = 20$   
iii) Plot Poisson pdf: use the following to graph for  $\lambda = 2, 5, 10, 15$  and  $n = 10, 20, 50$

`plot(0:n, dpois(0:n,lambda), type='h', xlab="occurences(arrival)",  
ylab="Probability")`

- iv) Plot Poisson cdf: use `ppois(0, n, lambda)` to plot  $F(x)$  for  $\lambda = 2, 5, 15$  and  $n = 10, 20, 50$

`plot(0:10,ppois(0:10,lambda),xlab="# arrivals", ylab="Cum Prob", type='s')`

- v) Generate vectors of  $n = 100, 500$  and  $5000$  realizations from  $\text{Pois}(\lambda = 5)$ .  
Compute the mean of each vector and compare to  $\lambda$ . Find the error magnitude for each simulation.

From R documentation:

**Usage**

`plot(x, y, ...)`

**Arguments**

- x** the coordinates of points in the plot. Alternatively, a single plotting structure, function or any *R* object with a plot method can be provided.  
**y** the y coordinates of points in the plot, *optional* if **x** is an appropriate structure.  
... Arguments to be passed to methods, such as [graphical parameters](#) (see [par](#)). Many methods will accept the following arguments:

**type**

what type of plot should be drawn. Possible types are

- "p" for **p**oints,

- "l" for lines,
- "b" for **both**,
- "c" for the lines part alone of "b",
- "o" for both 'overplotted',
- "h" for 'histogram' like (or 'high-density') vertical lines,
- "s" for stair steps,
- "S" for other steps, see 'Details' below,
- "n" for no plotting.

All other types give a warning or an error; using, e.g., type = "punkte" being equivalent to type = "p" for S compatibility. Note that some methods, e.g. [plot.factor](#), do not accept this.

main

an overall title for the plot: see [title](#).

sub

a sub title for the plot: see [title](#).

xlab

a title for the x axis: see [title](#).

ylab

a title for the y axis: see [title](#).

asp

the y/x aspect ratio, see [plot.window](#).

