UE2 - Christian Sallinger

(1) A left-turn lane problem

A civil engineer is studying a left-turn lane that is long enough to hold six cars. Let X be the number of cars in the lane at the end of a randomly chosen red light. The engineer believes that the probability that $X = x$ is proportional to $(x+1)(7-x)$.
(a) Find the probability mass function (pmf) of X.
(b) Compute the probability that X will be at least 4.
(c) Calulate the expectation and standard deviation of X. Note: R might be useful.

Solution

(a) Our random variable is obviously discrete and the probability mass function of a discrete random variable X is by definition given by

$$f_X(x) = P(X = x) = (x+1)(7-x) \cdot c$$

To find the value for c we note, that

$$\sum_{i=0}^{6} f_X(i) \stackrel{!}{=} 1$$

has to hold.

```
find_c <- function(x){(x+1)*(7-x)}
sum(find_c(0:6))
```

```
## [1] 84
```

This quick calculation shows, that $c = \frac{1}{84}$, so we get

$$f_X(x) = \frac{(x+1)(7-x)}{84}$$

(b) Here we just do a quick calculation.

```
f_X <- function(x){(x+1)*(7-x)/84}
sum(f_X(4:6))
```

```
## [1] 0.4047619
```

(c) We use our formulas and R to calculate the values.

```
mean_x = sum(c(0:6)*f_X(0:6))
standard_deviation = sqrt(sum(((c(0:6)-mean_x)^2)*f_X(0:6)))

mean_x
```

```
## [1] 3
```

```
standard_deviation
```

```
## [1] 1.732051
```

(2) Basketball free throws

Two professional basketball players, Tom and John, each throw ten free throws with a basketball. Tom makes 80% of the free throws he tries, while John makes 85% of the free throws he tries.
(a) What is the probability that the number of free throws that Tom will make is exactly 7
(b) What is the probability that the number of free throws that John will make is at least 8?
(c) Player who achives the highest score wins the game. It is assumed that the two players do not influence each other when throwing. What is the probability that neither Tom or John will win the game?
Hint: Use R-function dbinom() to calculate the probability mass functions.

Solution

(a) We use the R-function to calculate

$$P(X = 7)$$

```
dbinom(7,10,0.8)
```

```
## [1] 0.2013266
```

(b) Again we use the R-function to calculate

$$P(Y \geq 8) = P(Y = 10) + P(Y = 9) + P(Y = 8)$$

```
sum(dbinom(8:10,10,0.85))
```

```
## [1] 0.8201965
```

(c) Neither of them wins the game, if both make the same amount of free throws. So we have to calculate

$$\sum_{i=0}^{10} P(X = i) \cdot P(Y = i)$$

```
sum(dbinom(0:10,10,0.8)*dbinom(0:10,10,0.85))
```

```
## [1] 0.2276205
```

(3) Uniform-exponential relationship

(a) Let Y be an exponential random variable $Y \sim \exp(\lambda)$, i.e. its pdf is given by

$$f_Y(y) = \begin{cases} \lambda e^{-\lambda y}, & y \geq 0 \\ 0, & \text{else} \end{cases}$$

and its mean equals $\frac{1}{\lambda}$. Compute $P(Y > y)$.

(b) Let X be a random variable, uniformlly distributed on (0,1) . Find the cumulative distribution function of X. What is the distribution of $Z = -\log X$?

Solution

(a) We use our formula

$$P(Y > y) = 1 - P(Y \leq y) = 1 - \int_{-\infty}^{y} f_Y(t)dt = \begin{cases} 1, & y < 0 \\ e^{-\lambda y}, & y \geq 0 \end{cases}$$

(b) For a uniformlly distributed random variable X (on (0,1)) the pdf is of the form

$$f_X(x) = \begin{cases} 1, & x \in (0,1) \\ 0, & \text{else} \end{cases}$$

The cdf is then given by

$$F_X(x) = \begin{cases} 0, & x \leq 0 \\ x, & x \in (0,1) \\ 1, & \text{else} \end{cases}$$

Now to get the distribution of Z, we use a transformation. The transformation in question is $g : (0,\infty) \to (0,\infty)$ defined by

$$g(x) = -\log(x)$$

Our transformation is invertible with differentiable inverse $g^{-1}(x) = e^{-x}$. We now get the pdf of Z by

$$f_Z(z) = f_X(e^{-z})|-e^{-z}| = e^{-z}, \quad z > 0$$

(4) Hurricane insurance

An insurance company needs to asses the risk associated with providing hurricane insurance. During 22 years from 1990 through 2011 , Florida was hit by 27 major hurricanes (level 3 and above). The insurance company assumed Poisson distribution for modeling number of hurricanes.
(a) If hurricanes are independent and the mean has not changed, what is the probability of having a year in Florida with each of the following?
(1) No hits.
(2) Exactly one hit.
(3) More than two hits.
(b) Use R to estimate the number of hurricane hits that will occur with the probability 99.5%.
Hint: One of the following R-commands: dpois(), ppois(), qpois(), rpois() is applicable.

Solution

(a) From the lecture we know that $E(X) = \lambda$. According to the information we take $\lambda = \frac{27}{22}$. Now we use R to calculate the probabilities.

(1)

```
dpois(0,27/22)
```

```
## [1] 0.2930908
```

(2)

```
dpois(1,27/22)
```

```
## [1] 0.3597024
```

(3)

```
1-sum(dpois(0:2,27/22))
```

```
## [1] 0.1264803
```

(b) We calculate the value with the help of the quantile-function.

```
qpois(0.995,27/22)
```

```
## [1] 5
```

(5) Drug company

Manufacturing and selling drugs that claim to reduce an individual's cholesterol level is big business. A company would like to market their drug to women if their cholesterol is in the top 15%. Assume the cholesterol levels of adult American women can be desribed by a Normal model with a mean of 188mg/dL and a standard deviation of 24mg/dL.
(a) Use R to draw and label the Normal model.
(b) What percent of adult women do you expect to have cholesterol levels over 200mg/dL ?
(c) What percent of adult women do you expect to have cholesterol levels between 150mg/dL and 170mg/dL?
(d) Calculate the interquartile range of the cholesterol levels. Recall, the interquartile range is the diference between upper and lower quartile, i.e.

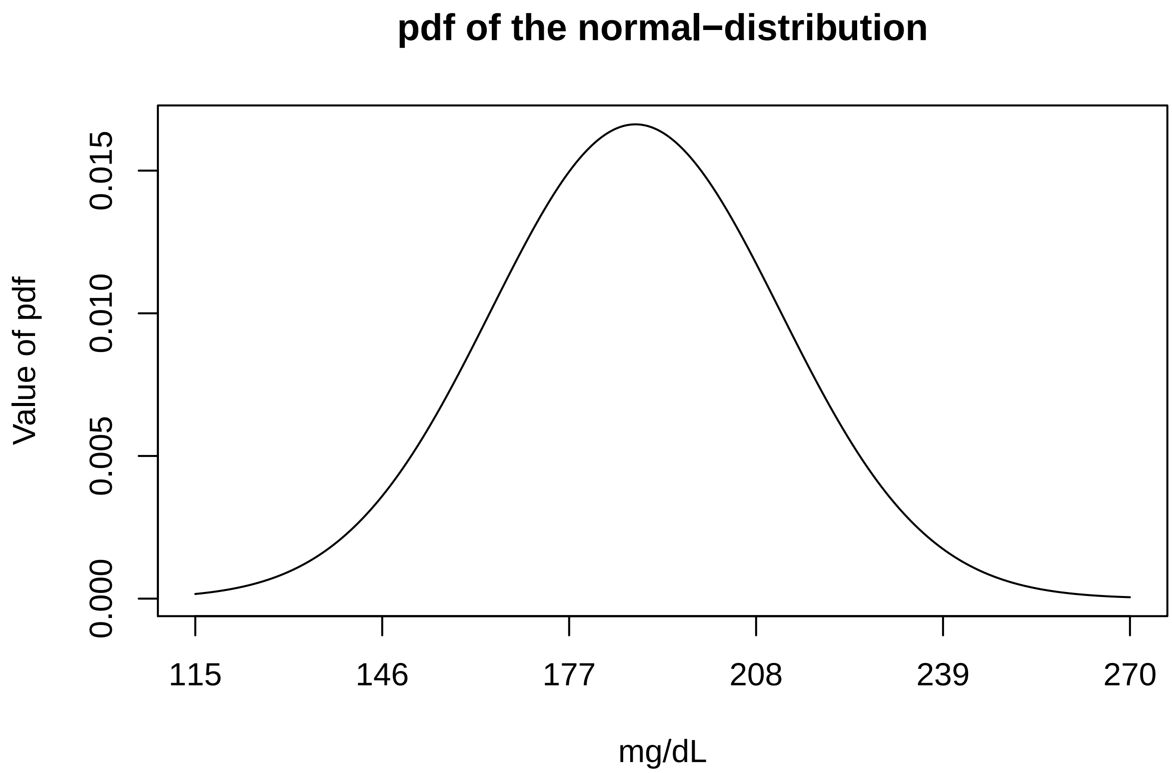
$$IQR = x_{0.75} - x_{0.25}$$

(e) Above what value are the highest 15% of women's cholesterol levels?
Hint: If using R for all computations the following commands pnorm(), qnorm() and dnorm() are useful. Otherwise values from Table of standard Normal distribution should be used.

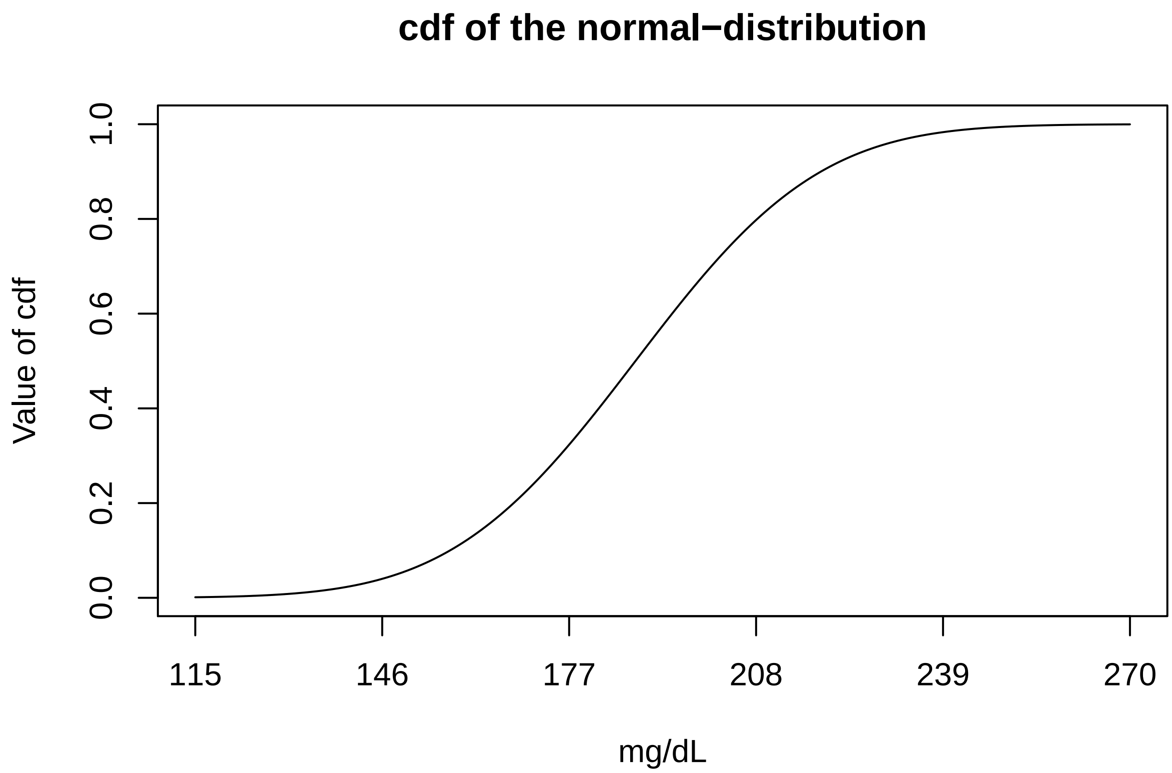
Solution

(a)

```
x <- seq(115,270,0.1)
y <- dnorm(x,188,24)
plot(x,y,type="l", main = "pdf of the normal-distribution",xlab = "mg/dL", ylab = "Value of pdf", xaxp
```




```
z <- pnorm(x,188,24)
plot(x,z,type = "l", main = "cdf of the normal-distribution",xlab = "mg/dL", ylab = "Value of cdf",xaxp
```



(b) We use R to calculate

$$P(X > 200) = 1 - P(X \leq 200).$$

```r
1 - pnorm(200,188,24)
```

```
## [1] 0.3085375
```

(c) We again use R to calculate

$$P(150 < X < 170) = P(X \leq 170) - P(X \leq 150)$$

```r
pnorm(170,188,24)-pnorm(150,188,24)
```

```
## [1] 0.1699546
```

(d) We use R.

```r
qnorm(0.75,188,24)-qnorm(0.25,188,24)
```

```
## [1] 32.37551
```

(e) Again R.

```r
qnorm(0.85,188,24)
```

```
## [1] 212.8744
```