# “Homework 2”

*Ballesteros, Mele*

# Part I

### 3.16

You are diagnosed with an uncommon disease. You know that there only is a 1% chance of getting it. Use the letter $D$ for the event “you have the disease” and $T$ for “the test says so.” It is known that the test is imperfect: $P(T|D) = 0.98$ and $P(T^c|D^c) = 0.95$:

a. Given that you test positive, what is the probability that you really have the disease?
b. You obtain a second opinion: an independent repetition of the test. You test positive again.Given this, what is the probability that you really have the disease?

---

**a.** We have to apply here Bayes Theorem:

$$P(D|T) = \frac{P(T|D)P(D)}{P(T|D)P(D) + P(T|D^c)P(D^c)} = \frac{0.98 \cdot 0.01}{0.98 \cdot 0.01 + 0.05 \cdot 0.99} = 0.17$$

**b.** We denote now $T_i$ the possitive result of test $i$, with $i = 1, 2$. Also we are using independence condition between both tests: $P(T_2 \cap T_1|D) = P(T_2|D)P(T_1|D)$.

Using Bayes Theorem again we gave:

$$\begin{aligned} P(D|T_2 \cap T_1) &= \frac{P(T_2 \cap T_1|D)P(D)}{P(T_2 \cap T_1|D)P(D) + P(T_2 \cap T_1|D^c)P(D^c)} = \\ &= \frac{P(T_2|D)P(T_1|D)P(D)}{P(T_1|D)P(T_2|D)P(D) + P(T_1|D^c)P(T_2|D^c)P(D^c)} = \\ &= \frac{0.98 \cdot 0.98 \cdot 0.01}{0.98 \cdot 0.98 \cdot 0.01 + 0.05 \cdot 0.05 \cdot 0.99} = 0.80 \end{aligned}$$

### 5.1

Let X be a continuous random variable with probability density function:

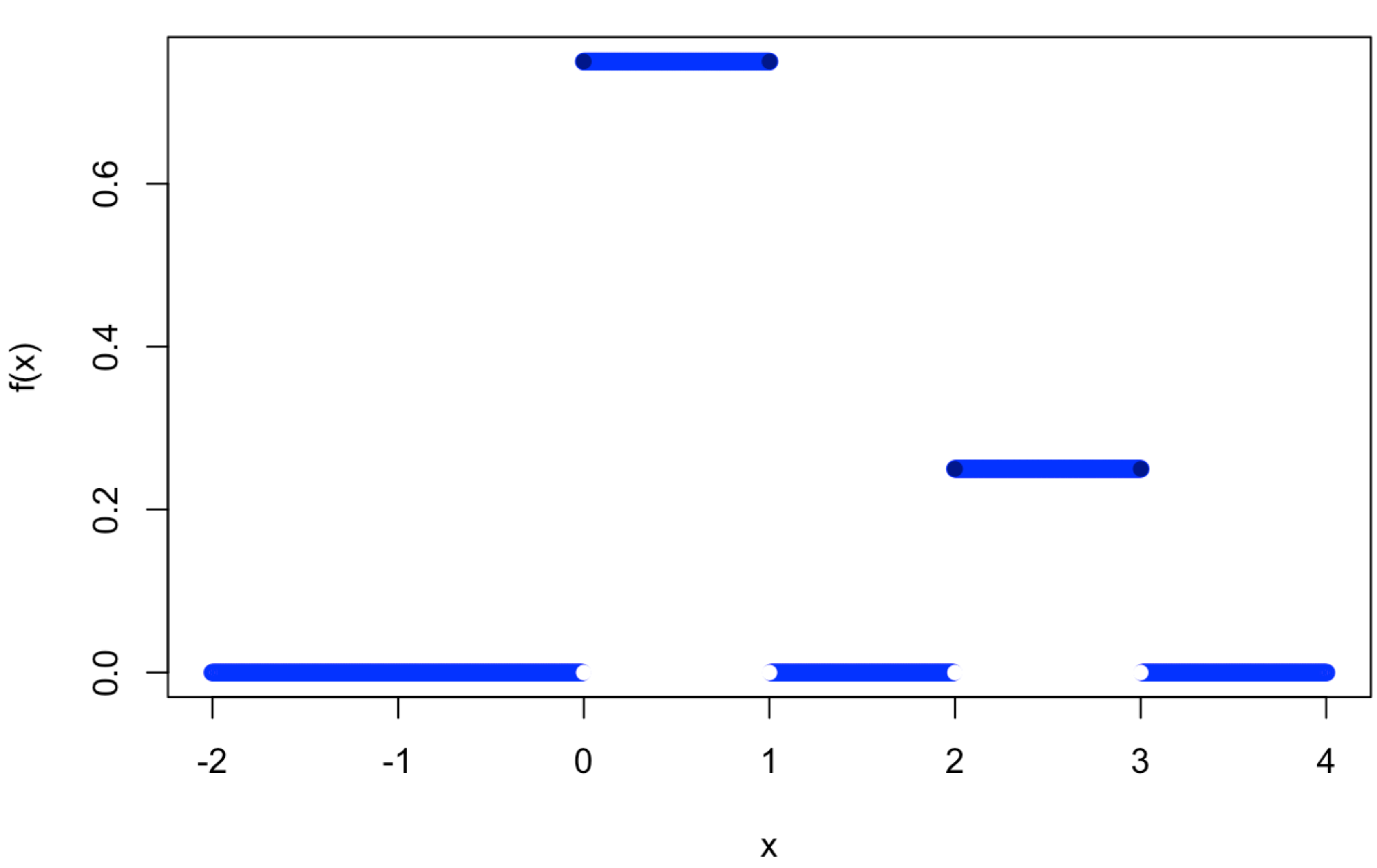
$$f(x) = \begin{cases} \frac{3}{4} & \text{for } \ 0 \leq x \leq 1 \\ \frac{1}{4} & \text{if } \ 2 \leq x \leq 3 \\ 0 & \text{elsewhere} \end{cases}$$

a. Draw the graph of f.
b. Determine the distribution function F of X, and draw its graph.

---

**a** Here there are the R commands to plot the $f(x)$ function

```
x= seq(-2,+4, by=0.01)
fx= (0<= x & x<=1) * 3/4 + (2<= x & x<= 3)* 1/4
plot(x,fx, ylab = 'f(x)', col='blue')
points(c(0,1,2,3),c(0,0,0,0),pch=16,bg='black', col='white')
points(c(0,1,2,3),c(3/4,3/4,1/4,1/4), pch=16, bg='black', col='blue4')
```

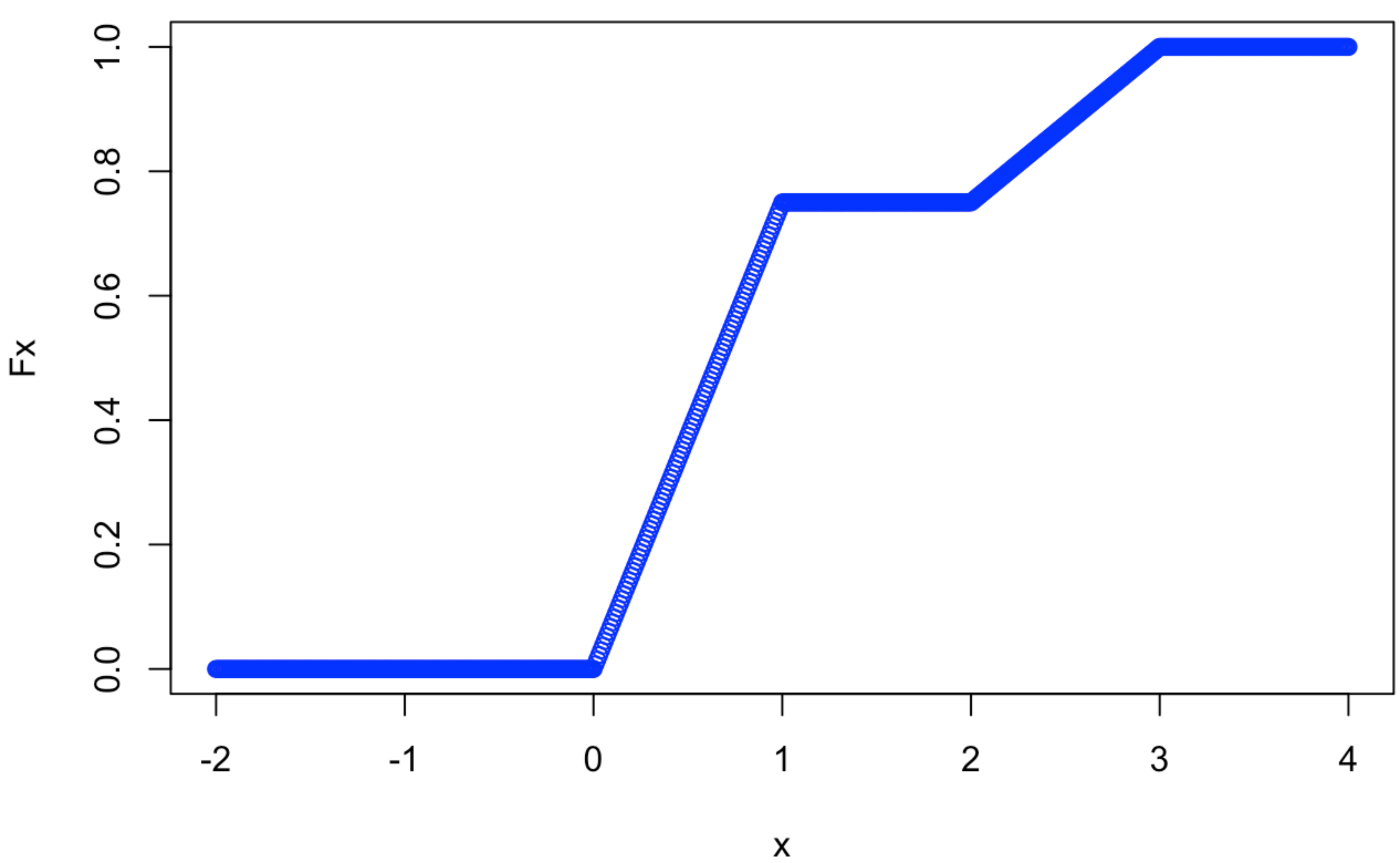



x not b

**b** considering that the CDF is: $F_X(x) = P(X \leq x)$ so for this case $F_X(x) = \int_{-\infty}^{x} f(x)dx$ , so:

$$F_X(x) = \begin{cases} 0 & \text{for } x < 0 \\ \frac{3b}{4} & \text{for } 0 \leq x < 1 \\ \frac{3}{4} & \text{for } 1 \leq x < 2 \\ \frac{b}{4} + \frac{1}{4} & \text{for } 2 \leq x < 3 \\ 1 & \text{if } x \geq 3 \end{cases}$$

```
x=seq(-2,4,by=0.01)
Fx=(x>=0 & x<1)*3*x/4 + (x>=1 & x<2)*3/4 + (x>=2 & x<3)*(1/4+ x/4)+ (x>=3)*1
plot(x,Fx, col='blue')
```




### 5.3

Let a continuous random variable $X$ be given that takes values in $[0, 1]$ , and whose distribution function $F$ :

$$F(x) = 2x^2 - x^4 \text{ for } 0 \leq x \leq 1$$

a. compute $P(\frac{1}{4} < x < \frac{3}{4})$
b. What is the probability density function of X ?

**a.** The probability asked is:

$P(\frac{1}{4} \leq X \leq \frac{3}{4}) = F(\frac{3}{4}) - F(\frac{1}{4}) = 2(\frac{3}{4})^2 - (\frac{3}{4})^4 - [2(\frac{1}{4})^2 - (\frac{1}{4})^4] = \frac{11}{16}$

**b.** To get the PDF we just have to derivate:

$$f(x) = \frac{d}{dx}F(x) = 4x - 4x^3$$

For $0 \leq x \leq 1$. For any other value of $x$, $f(x) = 0$

### 5.7

The score of a student on a certain exam is represented by a number between 0 and 1. Suppose that the student passes the exam if this number is at least $0.55$. Suppose we model this experiment by a continuous random variable S, the score, whose probability density function is given by

$$f(x) = \begin{cases} 4x & \text{for } 0 \leq x \leq \frac{1}{2} \\ 4 - 4x & \text{for } \frac{1}{2} \leq x \leq 1 \\ 0 & \text{elsewhere} \end{cases}$$

a. What is the probability that the student fails the exam?
b. What is the score that he will obtain with a 50% chance, in other words, what is the 50th percentile of the score distribution?

---

**a** The probability that a student fails the exam is $P(\text{failure}) = P(X < 0.55)$ so:

$$P(X < 0.55) = \int_{-\infty}^{0.55} f(x)dx = \int_{0}^{0.5} 4xdx + \int_{0.5}^{0.55} (4 - 4x)dx = 0.595$$

**b** The Quintile function is defined by $Q(p) = \inf\{x \text{ such that } F(x) = p\}$ so:

$$Q(0.5) = \inf\{x|F(x) = 0.5\} = \inf\{x|\int_{-\infty}^{x} f(x)dx\} = \inf\{x|\int_{0}^{x} 4xdx\} = \inf\{x|2x^2 = 0.5\}$$

$$x = 0.5$$

$$Q(0.5) = 0.5$$

### 5.11

Compute the median of an $Exp(\lambda)$ distribution.

---

The distribution is $f_X(x) = Exp(\lambda) = \lambda e^{-\lambda x}$ with a CDF $F_X(x) = 1 - e^{-\lambda x}$ if $0 \leq x$.

The median is the point where $F_X(x) = 0.5$, therefore we have to solve the following equation:

$0.5 = 1 - e^{-\lambda x}$

$x = \frac{ln(2)}{\lambda}$

### 5.14

Determine the '10-th' percentile of a standard normal distribution

---

The '10-th' percentile of a standard normal distribution is $\Phi(q_{0.1}) = 0.1$ and it can be found with the table of normal distribution's CDF:

$$\Phi(X \leq q_{0.1}) = 0.1 = 1 - 0.9 = 1 - \Phi(X \leq q_{0.9})$$

,but $\Phi(q_{0.9}) = \Phi(-q_{0.1})$ and $\Phi(q_{0.9}) = 1.28$ so:

$$q_{0.1} = -1.28$$

### 7.3

For a certain random variable $X$ it is known that $E[X] = 2$, $Var(X) = 3$. What is EX2?

---

The known equation for the variance is:

$Var(X) = E(X^2) - [E(X)]^2$

So just solving that for $E(X^2)$:

$E(X^2) = [E(X)]^2 + Var(X) = 2^2 + 3 = 7$

### 7.4

Let $X$ be a random variable with $E[X] = 2$, $Var(X) = 4$. Compute the expectation and variance of $(3 - 2X)$

---

For the change of units rule: $E(3 - 2X) = 3 - 2E(X) = -1$ and $Var(3 - 2X) = 2^2 Var(X) = 16$

### 7.7

Given is a random variable $X$ with probability density function $f$ given by $f(x) = 0$ for $x < 0$, and for x>1, and $f(x) = 4x - 4x^3$ for $0 \leq x \leq 1$. Determine the expectation and variance of the random variable 2X +3.

---

We can use the follogin formula to calculate the expectaton values:

$$E[g(x)] = \int g(x) f_X(x) dx$$

So we have then:

$$E(2x+3) = \int_0^1 (2x+3)(4x - 4x^3)dx = 4\int_0^1 (2x^2 - 2x^4 + 3x - 3x^3)$$
$$= 4(2/3 - 2/5 + 3/2 - 3/4) = 61/15$$

In order to calculate the variance, we can take a different approach. Lets firs calculate $E(x)$ and $E(x^2)$:

$$E(X) = \int_0^1 xf(x)dx = \int_0^1 x(4x - 4x^3)dx = 8/15$$

$$E(X^2) = \int_0^1 x^2 f(x)dx = \int_0^1 x^2(4x - 4x^3)dx = 1/3$$

$$Var(X) = E(X^2) - E^2(X) = 11/225$$

Now we van apply the following property: $Var(rX + s) = r^2 Var(X)$

Which then leds to the following result: $Var(2x + 3) = 2^2 Var(x) = 44/225$

### 9.1

The joint probabilities $P(X = a, Y = b)$ of discrete random variables $X$ and $Y$ are given in the following table (which is based on the magical square in Albrecht D?urer's engraving *Melencolia I* ). Determine the marginal probability distributions of $X$ and $Y$ ,i.e.,determine the probabilities $P(X = a)$ and $P(Y = b)$ for a,b =1,2,3,4.

| b | a1 | a2 | a3 | a4 |
|---|---|---|---|---|
| 1 | 16/136 | 3/136 | 2/136 | 13/136 |
| 2 | 5/136 | 10/136 | 11/136 | 8/136 |
| 3 | 9/136 | 6/136 | 7/136 | 12/136 |
| 4 | 4/136 | 15/136 | 14/136 | 1/136 |

---

The marginal probability is $P(X = a) = P(X = a, Y = \omega) = \sum_{b=1}^{4} P(X = a, Y = b) = \frac{1}{4}$ for $a \in [1, 2, 3, 4]$

and $P(Y = b) = P(Y = b, X = \omega) = \sum_{a=1}^{4} P(Y = b, X = a) = \frac{1}{4}$ for $b \in [1, 2, 3, 4]$

### 9.3

Let $X$ and $Y$ be two random variables, with joint distribution the Melencolia distribution, given by the table in Exercise 9.1. What is

a. $P(X = Y)$?
b. $P(X + Y = 5)$?
c. $P(1 < X \leq 3, 1 < Y \leq 3)$?
d. $P((X, Y) \in (1, 4) \text{ x } (1, 4))$?

---

**a.** $P(X = Y) = \sum_{i=1}^{4} P(X = i, Y = i) = 16/136 + 10/136 + 7/136 + 1/136 = 1/4$

**b.**
$P(X + Y = 5) = P(X = 1, Y = 4) + P(X = 4, Y = 1) + P(X = 2, Y = 3) + P(X = 3, Y = 2) = 4/136 + 13/136 + 6/136 + 11/136 = 1/4$

**c.**
$P(1 < X \leq 3, 1 < X \leq 3) = P(X = 2, Y = 2) + P(X = 2, Y = 3) + P(X = 3, Y = 2) + P(X = 3, Y = 3) = 10/136 + 6/136 + 11/136 + 7/136 =$

**d.**
$P((X, Y) \exists \{1, 4\} \times \{1, 4\} = P(X = 1, Y = 1) + P(X = 1, Y = 4) + P(X = 4, Y = 1) + P(X = 4, Y = 4) = 16/136 + 13/136 + 4/136 + 1/136 =$

### 9.7

To investigate the relation between hair color and eye color, the hair color and eye color of 5383 persons was recorded. The data are given in the following table:

| | Hair color | | |
|---|---|---|---|
| Eye color | Fair / red | Medium | Dark /Black |
| Light | 1168 | 825 | 305 |
| Dark | 573 | 1312 | 1200 |

Eye color is encoded by the values 1 (Light) and 2 (Dark), and hair color by 1 (Fair/red), 2 (Medium), and 3 (Dark/black). By dividing the numbers in the table by $5383$, the table is turned into a joint probability distribution for random variables $X$ (hair color) taking values 1 to 3 and $Y$ (eye color) taking values 1 and 2.

a. Determine the joint and marginal probability distributions of $X$ and $Y$.
b. Find out whether $X$ and $Y$ are dependent or independent.

---

**a.** The joint probability distribution of $X$ and $Y$ is

| **b** | **a** | | | |
|---|---|---|---|---|
| | 1 | 2 | 3 | P(Y= b) |
| 1 | 0.22 | 0.15 | 0.06 | 0.43 |
| 2 | 0.11 | 0.24 | 0.22 | 0.57 |
| P(X=a) | 0.33 | 0.39 | 0.28 | 1 |

**b.** Since $P(X=1, Y=1) = \frac{1168}{5383} \neq \frac{1741}{5383}\frac{2298}{5383} = P(X=1)P(Y=1)$

we answer that $X$ and $Y$ are dependent.

### 9.12

The joint probability density function f of the pair $(X, Y)$ is given by $f(x, y) = K(3x^2 + 8xy)$ for $0 \leq x \leq 1$ and $0 \leq y \leq 2$, and $f(x, y) = 0$ for all other values of $x$ and $y$. Here $K$ is some positive constant.

a. Find $K$.
b. Determine the probability $P(2X \leq Y)$.

**a.** To find $K$ we must use the condition that the PDF has to be normalized:

$\int_0^2 \int_0^1 K(3x^2 + 8xy)dxdy = 10K = 1$

$K = 1/10$

**b.**

$P(2X \leq Y) = P(X \leq Y/2) = \int_0^2 \int_0^{y/2} \frac{1}{10}(3x^2 + 8xy)dxdy = \frac{1}{10}\int_0^2 (y^3/8 + y^3) = 9/20$

### 10.8

Let $X$ and $Y$ be random variables such that $E[X] = 2$ , $E[Y] = 3$ , and $Var(X) = 4$

a. Show that $E[X^2] = 8$.
b. Determine the expectation of $-2X^2 + Y$.

---

**a.** Using the know equality $Var(X) = E(X^2) - E(X)^2$ we have:

$$E(X^2) = Var(X) + E(X)^2 = 4 + 2^2 = 8$$

**b.** We can use here the this property: $E(rX + sY) = rE(X) + sE(Y)$:

$$E(-2X^2 + Y) = -2E(X^2) + E(Y) = -28 + 3 = -13$$

# Part II

## 1

Suppose that X has a Unif(0,1) distribution. Construct a random variable $Y = g(X)$ for some deterministic function $g(?)$ such that its $pdf$ will be:

$$f_Y(y) = \begin{cases} \frac{3}{8}y^2 & \text{for } y \in (0,2) \\ 0 & \text{otherwise} \end{cases}$$

---

The uniform distribution of X has the following CDF:

$$F_X(x) = \begin{cases} 0 & \text{if } x < 0 \\ x & \text{if } 0 \leq x \leq 1 \\ 1 & \text{if } 1 < x \end{cases}$$

Now we calculate the CDF of Y:

$$F_Y(y) = \int f_Y(y)dy = \int \frac{3}{8}y^2\,dy = \begin{cases} 0 & \text{if } y < 0 \\ \frac{1}{8}y^3 & \text{if } 0 \leq y \leq 2 \\ 1 & \text{if } 2 < y \end{cases}$$

Now we can develop the following equality:

$$F_Y(f) = P(Y \leq y) = P(g(X) \leq y) = P(X \leq g^{-1}(y)) = F_X(g^{-1}(y)) = g^{-1}(y) = \frac{1}{8}y^3$$

We find now the inverse of the inverse function in the last equality and find that:

$$g(y) = 2y^{1/3}$$

Therefore the solution is:

$$Y = g(X) = 2X^{1/3}$$

# 2

In a die-coin experiment, a fair die is rolled and then a fair coin is tossed the number of times showing on the die. Let $N$ denote the die score and $X$ the number of heads. Find $E(X)$ and $Var(X)$.

---

In the dice-coin experiment, lets first talk about the coin toss. The PMF is simply the binomial distribution. Then the expected value $E_n(X)$ of success $x$ (which denotes de number of heads) in $n$ tosses, is given by:

$$(x) = \sum_{x=1}^{n} xP(X = x) = \sum_{x=1}^{n} x\frac{1}{2^n}\binom{n}{x}$$

Now considering the first part of the experiment, the dice, we want to compute the expected value of that expectec value. Since the dice is fair, the probability of each $E_n(x)$ is $1/6$

$$E(E_n(x)) = \sum_{n=1}^{6} E_n(x)P(E_n(x)) = \sum_{n=1}^{6}\sum_{x=1}^{n} x\frac{1}{2^n}\binom{n}{x}\frac{1}{6} = E(x)$$

Also we will need $E(E_n(x^2))$ to calculate the $var(x)$:

$$E(E_n(x^2)) = \sum_{n=1}^{6}\sum_{x=1}^{n} x^2\frac{1}{2^n}\binom{n}{x}\frac{1}{6} = E(x^2)$$

And the variance is of course:

$$Var(x) = E(x^2) - [E(x)]^2$$

We can compute this sums in R as follows:

```
# Compute expected value of x
sum = 0
for (n in 1:6) {
  for (x in 1:n) {
    sum = sum + x*dbinom(x, size=n, prob=0.5)/6
  }
}

# Compute expected value of x^2
sum2 = 0
for (n in 1:6) {
  for (x in 1:n) {
    sum2 = sum2 + x*x*dbinom(x, size=n, prob=0.5)/6
  }
}
```

Then the result is that $E(x) = 1.75$:

```
sum
```

```
## [1] 1.75
```

And $Var(x) = 1.60$:

```
sum2-sum*sum
```

```
## [1] 1.604167
```