Task 1

Assume that continuous random variable X has PDF that is non-zero only on segment $[a, b]$ and strictly positive on open interval (a, b). Median of random variable X is value $m \in (a, b)$ such that $P(X < m) = P(X > m) = 1/2$.

Prove that for symmetrical PDFs median is equal to expected value.

Provide an example of PDF such that median is larger than the expected value.

Let X be random variable, f be strictly increasing function and $Y = f(X)$. What can you say about medians of X and Y?

• As PDF_X is symmetrical on segment $[a, b]$, it's symmetric about the middle value of this segment $\frac{a+b}{2}$.

$P(X < m) = P(X > m) = 1/2$ means that half of the values of X are less than m and the other half is greater than m. This actually means that values of X are distributed symmetrically about m, so PDF_X is symmetric about m. Hence $m = \frac{a+b}{2}$.

Let's consider continuous random variable $Y = X - m$. We can say that PDF_Y is similar to PDF_X: it's non-zero on segment $\left[\frac{a-b}{2}, \frac{b-a}{2}\right]$ and is symmetric about 0.

From the symmetry of PDF_Y about 0 we can conclude that random variable $-Y$ has the same distribution, hence $E(-Y) = EY$.

On the other hand, because of properties of expected value, $E(-Y) = -EY$.

As such we get that $EY = -EY$, so $EY = 0$.

Now we can use properties of expected value again to find EX:

$EY = 0$

$E(X - m) = 0$

$EX - E(m) = 0$, as m is a constant $E(m) = m$, so

$EX - m = 0$

$EX = m$, q.e.d.

• We can take some random variable with non-symmetric distribution and increasing pdf, for example let X have $PDF(x) = \begin{cases} 2x, & for\ x \in [0; 1] \\ 0, & otherwise \end{cases}$.

Then $EX = \int_0^1 (x \cdot 2x)dx = \frac{2}{3}x^3|_0^1 = \frac{2}{3} \approx 0.67$.

We can also calculate the median.

$$P(X < m) = \int_0^m 2x dx = x^2|_0^m = m^2$$

$$m^2 = \frac{1}{2} \to m = \frac{\sqrt{2}}{2} \approx 0.71$$

For such random variable median is greater then EX.

- If function f is strictly increasing, then it basically means that

for any a, b if $a < b$ then $f(a) < f(b)$.

Then we can say that for any $X < m$ we get $f(X) < f(m)$, or in other terms $Y < f(m)$.

Taking that into account we can see that $P(X < m)$ is equal to $P\big(Y < f(m)\big)$ and $P(X > m) = P\big(Y > f(m)\big)$. So if m is the median of X, then both of these probabilities are equal 1/2, and as such $f(m)$ is the median of Y.