SGA - Probability Theory - Isupov Ilya

March 2024

Telegram: @Jeremix

1 The first task

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(a,b) such that P(X<m)=P(X>m)=1/2.

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

Let's show that E(X) = m:

The formula of the expected value E(X) of a continuous random variable:

$$EX = \int_{-\infty}^{\infty} x \cdot p(x) \cdot dx$$

In our case we know that continuous random variable X has PDF that is non-zero only on segment [a,b], so the formula of the expected value:

$$EX = \int_{a}^{b} x \cdot p(x) \cdot dx$$

p(x) - the PDF of X.

Since the PDF is symmetrical, for all x in [a,b]:

$$p(x) = p(b + a - x)$$

$$EX = \int_{a}^{b} xp(x)dx = \int_{a}^{b} xp(b + a - x)dx.$$

Let's make a replacement:

$$t = b + a - x$$

$$x = b + a - t$$

$$EX = \int_{a}^{b} (b + a - t)p(t)(-dt) = b + a - \int_{a}^{b} tp(t)dt.$$

Since the integral of $t \cdot p(t)$ by [a, b] is the expected value of a random variable $b + a - x$, which is also symmetric, it should be equal to $b + a - EX$.

Thus:

$$EX = (b + a) - EX$$

$$EX = \frac{(a + b)}{2}$$

Let's find the median:

$$p(x) = (a + b + x)$$

$$\int_a^b p(x)dx = 1$$

$$\int_a^{\frac{a+b}{2}} p(x)dx + \int_{\frac{a+b}{2}}^b p(x)dx = \int_a^{\frac{a+b}{2}} p(x)dx + \int_{\frac{a+b}{2}}^b p(a+b-x)dx$$

$$u = a + b - x$$

$$\int_a^{\frac{a+b}{2}} p(x)dx + \int_a^{\frac{a+b}{2}} p(u)du = 1$$

$$\int_a^{\frac{a+b}{2}} p(x)dx = \int_a^{\frac{a+b}{2}} p(u)du$$

$$2\int_a^{\frac{a+b}{2}} p(x)dx = 1$$

$$\int_a^{\frac{a+b}{2}} p(x)dx = \frac{1}{2}$$

Thus $\frac{a+b}{2}$ is the median.

$$m = \frac{a+b}{2} = EX$$

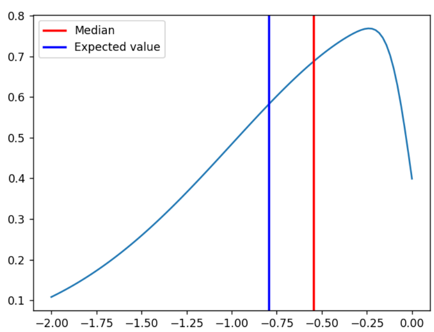
Provide an example of PDF such that median is larger than the expected value.
One of the examples where the median is larger than the expected value is a left skewed distribution. Let's implement it using Python and show left-skewed distribution graph:

```
data = skewnorm.rvs(a=-10, size=100)
median = np.median(data)
EX = skewnorm.mean(a=-10)
print(f'median = {round(median, 2)}')
print(f'Expected value = {round(EX, 2)}')

x = np.linspace(-2,  stop: 0,  num: 100)
plt.plot( *args: x, skewnorm(-10).pdf(x))
plt.axvline(median, color='r', linewidth=2, label='Median')
plt.axvline(EX, color='b', linewidth=2, label='Expected value')
plt.legend()

week_5 ×

C:\Users\y2966\PycharmProjects\Probability_theory\.venv\Scripts\pyth
median = -0.55
Expected value = -0.79
```



Other example:

$$PDF_X(x) = x \text{ if } 0 < x < \sqrt{2}$$
$$PDF_X(x) = 0 \text{ otherwise}$$

Median:

$$\int_0^m x dx = \frac{1}{2}$$
$$m = 1$$

Expected value:

$$EX = \int_0^{\sqrt{2}} x \cdot x \cdot dx = 0.94$$

In this case we have median = 1 and EX = 0.94. Thus, the 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?

The median of X will be converted to the median of Y using the function f. Thus, if m_X is the median of X, then $f(m_X)$ will be the median of Y since f is strictly increasing function.

$$Y = f(X)$$
$$P(X > m_X) = P(X < m_X) = \frac{1}{2}$$
$$P(Y > m_Y) = P(Y < m_Y) = \frac{1}{2}$$

Since f is strictly increasing function, f saves the order:

$$P(f(X) > f(m_X)) = P(f(X) < f(m_X)) = \frac{1}{2}$$

$$P(X > m_X) = P(f(X) > f(m_X)) = P(Y > f(m_X)) = \frac{1}{2} = P(Y > m_Y)$$
$$f(m_X) = m_Y$$

The median of Y will be the result of applying the function f to the median of X.

2 The second task

Let X be uniform random variable on a segment [0,2]. Consider random variable $Y = X^2$. Find CDF and PDF of Y. Is PDF a bounded function?

Let's find the PDF of X:

$$PDF_X(x) = \frac{1}{2} \text{ if } 0 < x < 2$$
$$PDF_X(x) = 0 \text{ otherwise}$$

Let's find the CDF of X:

$$CDF_X(x) = \int \frac{1}{2} dx = \frac{1}{2} x \text{ if } 0 \leq x < 2$$

$$CDF_X(x) = 0 \text{ if } x < 0$$

$$CDF_X(x) = 1 \text{ if } 2 \leq x$$

Let's find the CDF of Y:

$$CDF_Y(y) = F_Y(y) = P(Y \leq y) = P(X^2 \leq y) = P(X \leq \sqrt{y}) = CDF_X(\sqrt{y}) = \frac{\sqrt{y}}{2}$$

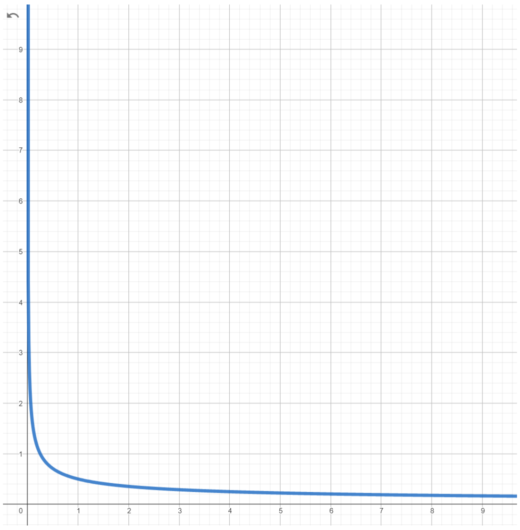
$$CDF_Y(y) = \frac{\sqrt{y}}{2} \text{ if } 0 < \sqrt{y} < 2$$
$$CDF_Y(y) = 0 \text{ if } y < 0$$
$$CDF_Y(y) = 1 \text{ if } y > 0$$

Let's find the PDF of Y by means of differentiation:

$$PDF_Y(y) = (\sqrt{y})' = \frac{1}{4 \cdot \sqrt{y}} \text{ if } 0 < y < 4$$

$$PDF_Y(y) = 0 \text{ otherwise}$$

PDF of Y isn't bounded function because if we increase y then PDF will be tends to zero, if we decrease y then PDF will be tends to infinity. We can draw this graph:



3 The third task

A fair coin is tossed 400 times. Let X be number of heads. Prove that $P(X > 240) \leq \frac{1}{32}$. The probability of an eagle falling on one roll (if it's a fair coin) is equal to:

$$p = \frac{1}{2}$$

Let's find the expected value and variance of X:

$$EX = n \cdot p = 400 \cdot \frac{1}{2} = 200$$

$$Var(X) = n \cdot p \cdot (1 - p) = 400 \cdot 0.5(1 - 0.5) = 100$$

Let's use Chebyshev's rule and find $P(|X - 200| > 40)$:

$$P(|X - 200| > 40) \leq \frac{100}{40^2} = \frac{1}{16}$$

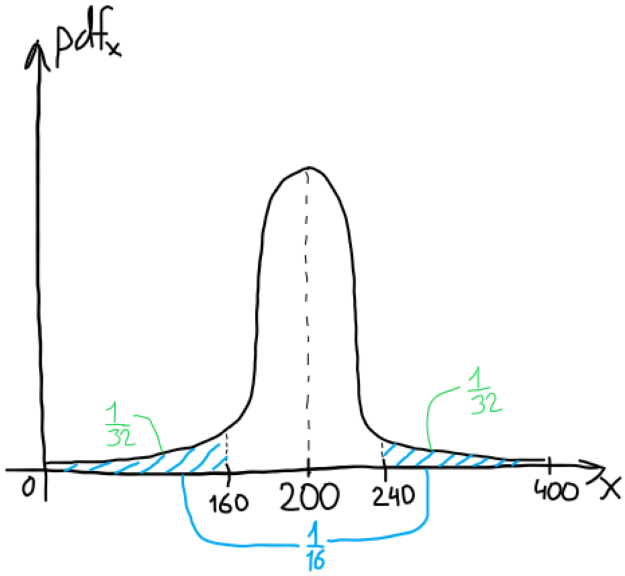
$P(|X - 200| > 40)$ means the probability of deviation from the expected value by more than 40.
$P(|X - 200| > 40)$ includes $P(X > 240)$ and $P(X < 160)$. We know that distribution of X is symmetric, it means that:

$$P(X > 240) = P(X < 160)$$

$$P(|X - 200| > 40) = P(X > 240) + P(X < 160) \leq \frac{1}{16}$$

Thus:

$$P(X > 240) \leq \frac{1}{16} \cdot \frac{1}{2} \leq \frac{1}{32}$$



4 The fourth task

Let X and Y be two independent normally distributed random variables with expected value 0 and variance 1. Find their joint PDF. Plot its level curves.
Let's write the formula of joint PDF of two independent random variables:

$$PDF_{XY}(x, y) = PDF_x(x) \cdot PDF_Y(y)$$

Let's find $PDF_x(x)$ and $PDF_Y(y)$ knowing that X and Y are normally distributed random variables:

$$PDF_X(x) = \frac{1}{\sigma \cdot \sqrt{2\pi}} \cdot e^{-\frac{1}{2}(\frac{x-\mu}{\sigma})^2} = \frac{1}{1 \cdot \sqrt{2\pi}} \cdot e^{-\frac{1}{2}(\frac{x-0}{1})^2} = \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{x^2}{2}}$$

$$PDF_Y(y) = \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{y^2}{2}}$$

Let's write the joint PDF of two independent random variables X and Y:

$$PDF_{XY}(x,y) = \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{x^2}{2}} \cdot \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{y^2}{2}} = \frac{1}{2\pi} \cdot e^{-\frac{x^2}{2}-\frac{y^2}{2}} = \frac{1}{2\pi} \cdot e^{-\frac{x^2+y^2}{2}}$$

Let's denote $PDF_{XY}(x,y)$ as Z and write the equation in a different form:

$$PDF_{XY}(x,y) = Z$$
$$Z = \frac{1}{2\pi} \cdot e^{-\frac{x^2+y^2}{2}}$$
$$2\pi \cdot Z = e^{-\frac{x^2+y^2}{2}}$$
$$\ln(2\pi \cdot Z) = -\frac{x^2+y^2}{2}$$
$$x^2 + y^2 = -2 \cdot \ln(2\pi \cdot Z)$$

As we can see on the left side of our equation, we have the equation of a circle. This means that $PDF_XY(x,y)$ will take equal values on the circle.
Let's build a graph for several values of X and Y: