Continuous RV: PDF and CDF

Online S520 Supplemental Material

An example of continuous case: pdf, cdf, find probability, median and mean.

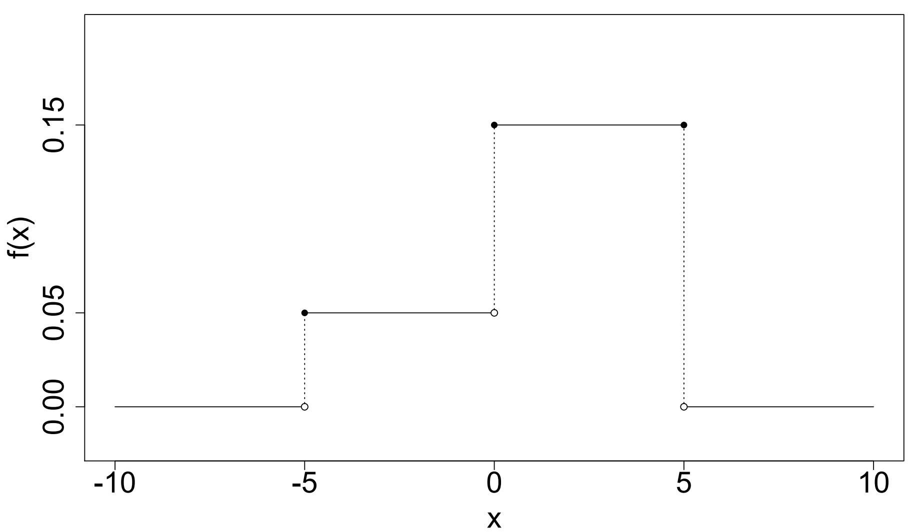
Let X be a continuous random variable with the following probability density function (PDF):

$$f(x) = \begin{cases} 0.05 & -5 \le x < 0 \\ 0.15 & 0 \le x \le 5 \\ 0 & \text{otherwise.} \end{cases}$$

1. Make a graph of the PDF, f.
2. Verify $f(x)$ is a pdf.
3. Find $F(y)$, the cumulative distribution function of X, and make a graph of $F(x)$.
4. Find probabilities Find $P(|X| < 2.5)$.
5. Find the median.
6. Compare the median with mean/expected value.

Solutions:

1. Note: The points at (-5, 0.05), (0, 0.15) and (5, 0.15) are closed.

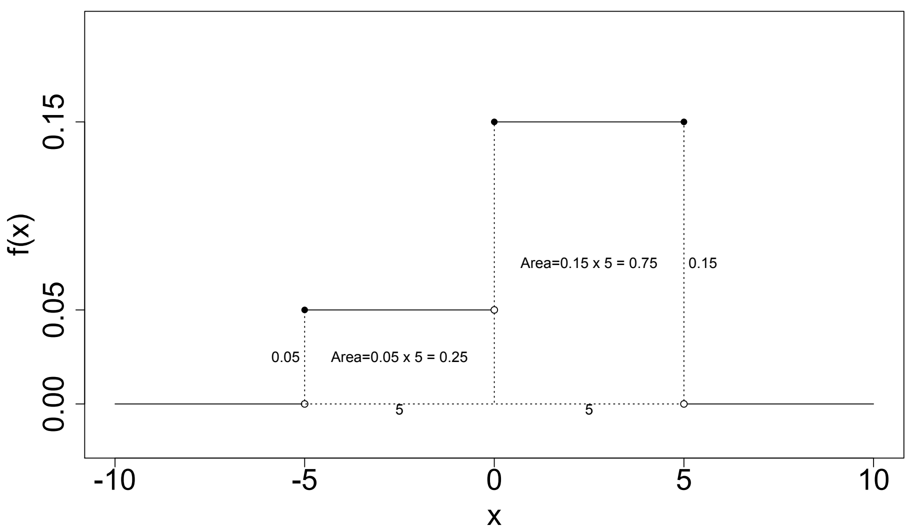


2. To check that a function is a pdf, we need check the following two conditions:

   - $f(x) \geq 0$ for every $x \in \Re$.
   - $\text{Area}_{(-\infty,\infty)}(f) = \int_{-\infty}^{\infty} f(x)\,dx = 1$.

   The first one is easy to check from the observation of $f(x)$. To check the second one, you can either use geometric calculation or calculus.

   [Use geometry] See graph below. The area under the graph of f and above the horizontal axis is the sum of the areas of two rectangles, One has sides 5 and 0.05, and thus the area is $5 \times 0.05 = 0.25$. The other one has sides 5 and 0.15, and the area is $5 \times 0.15 = 0.75$. So the total area is $0.25 + 0.75 = 1$.



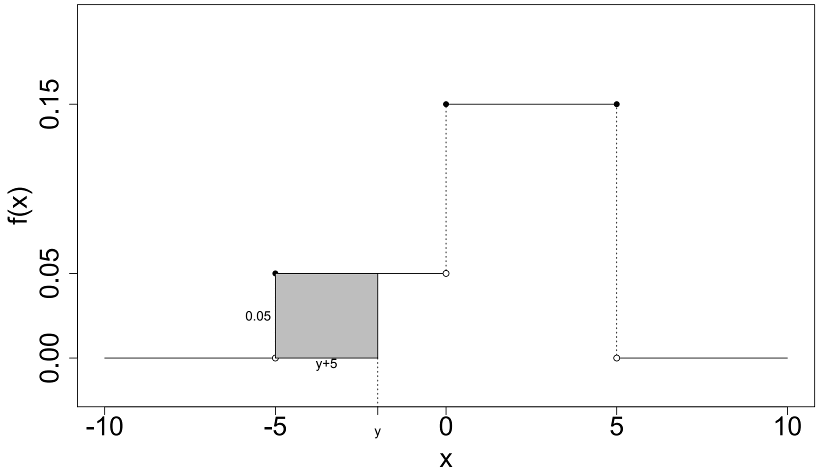
   [Use calculus]

$$\begin{aligned}
\int_{-\infty}^{\infty} f(x)\,dx &= \int_{-5}^{0} 0.05\,dx + \int_{0}^{5} 0.15\,dx \\
&= 0.05x|_{-5}^{0} + 0.15x|_{0}^{5} \\
&= [0.05 \cdot 0 - 0.05 \cdot (-5)] + [0.15 \cdot 5 - 0.15 \cdot 0] \\
&= 1
\end{aligned}$$

3. To derive the cdf, we first observe that the possible values of X are $X(S) = [-5, 5]$. So it is evident that $F(y) = 0$ if $y < -5$ and $F(y) = 1$ if $y > 5$.

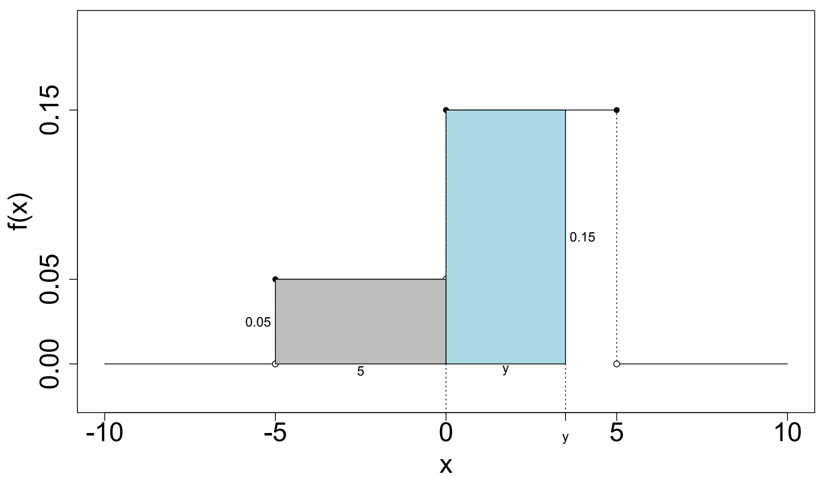
   [Use geometry] To find $F(y)$ for $y \in [-5, 5]$, we need calculate the area of the region bounded by f and the horizontal axis and between -5 and y.

   If $y \in [-5, 0)$, this region is a rectangle with base $y + 5$ and height 0.05. So the area is $0.05(y + 5)$.



   Hence,
$$F(y) = 0.05(y + 5), \quad \text{for } y \in [-5, 0)$$

   If $y \in [0, 5]$, this region can be divided into two rectangles. One has base 5 and height 0.05 and the other one has base y and height 0.15. So the area is $0.05 \cdot 5 + 0.15y$.



Hence,

$$F(y) = 0.25 + 0.15y, \quad \text{for } y \in [0,5]$$

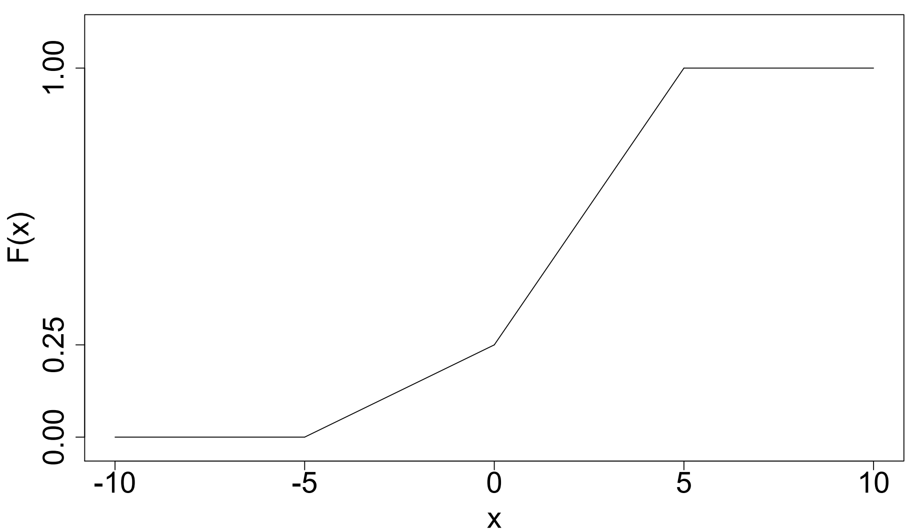
[Use calculus]

$$F(y) = \int_{-\infty}^{y} f(x)\,dx = \begin{cases} \int_{-\infty}^{y} f(x)\,dx & y < -5 \\ \int_{-\infty}^{y} f(x)\,dx & -5 \le y < 0 \\ \int_{-\infty}^{y} f(x)\,dx & 0 \le y \le 5 \\ \int_{-\infty}^{y} f(x)\,dx & y > 5 \end{cases}$$

$$= \begin{cases} \int_{-\infty}^{y} 0\,dx & y < -5 \\ \int_{-\infty}^{-5} 0\,dx + \int_{-5}^{y} 0.05\,dx & -5 \le y < 0 \\ \int_{-\infty}^{-5} 0\,dx + \int_{-5}^{0} 0.05\,dx + \int_{0}^{y} 0.15\,dx & 0 \le y \le 5 \\ \int_{-\infty}^{-5} 0\,dx + \int_{-5}^{0} 0.05\,dx + \int_{0}^{5} 0.15\,dx + \int_{5}^{y} 0\,dx & y > 5 \end{cases}$$

$$= \begin{cases} 0 & y < -5 \\ 0.05(y+5) & -5 \le y < 0 \\ 0.25 + 0.15y & 0 \le y \le 5 \\ 1 & y > 5 \end{cases}$$

Graph of $F(x)$:



4.

$$P(|X| < 2.5) = P(-2.5 < X < 2.5)$$

If $F(x)$ is available, we can just plug in numbers in the CDF to calculate the probability. If we don't have $F(x)$ ready to use, we can also use geometric calculation.

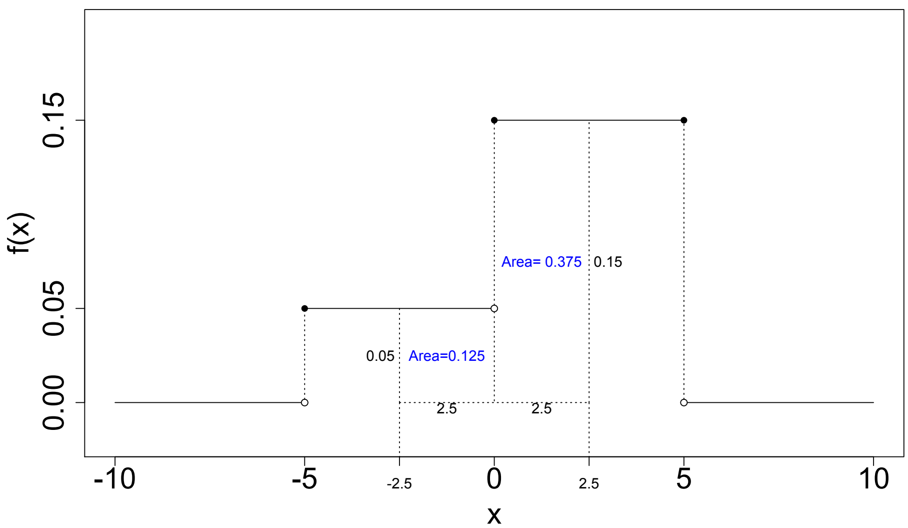
[Use $F(x)$]

$P(|X| < 2.5) = F(2.5) - F(-2.5) = [0.25 + 0.15 \cdot 2.5] - [0.05(-2.5 + 5)] = 0.625 - 0.125 = 0.5.$

[Use geometry]

$$P(|X| < 2.5) = \text{Area}_{(-2.5,2.5)}(f)$$

The area of this region is the sum of areas of two rectangles with base 2.5 and heights 0.05 and 0.15 respectively. So the total area is $2.5 \cdot 0.05 + 2.5 \cdot 0.15 = 0.5$. (See graph below)



Specifically for this question, we can also use the property of symmetry. The area between -2.5 and 2.5 is the same as the sum of the areas below -2.5 and above 2.5. So

$$P(|X| < 2.5) = P(|X| > 2.5) = 1/2.$$

5. Median, or the second quartile, q_2, is the 0.5 quantile of a random variable. We need find a value q_2 such that $P(X < q_2) = 0.5$.

   If $F(x)$ is available, we can just plug in numbers in the CDF to solve for q_2. If we don't have $F(x)$ ready to use, we can also use geometric calculation.

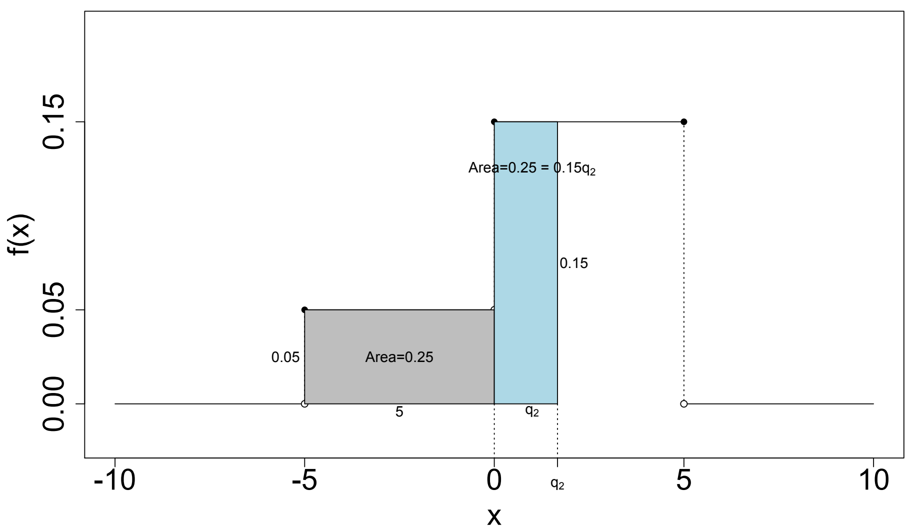
   [Use $F(x)$] q_2 is a value in [-5, 5] for which $F(q_2) = 0.5$.

   If q_2 is in [-5, 0), $F(q_2) = 0.05(q_2 + 5)$ ranges from 0.05(-5+5)=0 (inclusive) to 0.05(0+5) = 0.25 (exclusive). So it is not possible for q_2 to lie in [-5, 0).

   If q_2 is in [0, 5], $F(q_2) = 0.25 + 0.15q_2$.

$$0.5 = 0.25 + 0.15q_2 \rightarrow q_2 = \frac{0.25}{0.15} = \frac{5}{3}.$$

   [Use geometry] The area of a region between -5 and q_2 under the graph of f and above the horizontal axis is 0.5. Since the region between -5 and 0 is a rectangle with area 0.25, q_2 must lie on the interval (0, 5]. The region between 0 and q_2 is also a rectangle with area 0.25. The base is q_2 and the height is 0.15, so $0.15q_2 = 0.25$ gives $q_2 = 5/3$.



6. To compare the median with the expected value, we can calculate expected value either with calculus or property of expected values) and then compare two numbers.

[Use calculus to find the expected value]

$$\begin{aligned}
\mu = E(X) = \int_{-\infty}^{\infty} x f(x)\,dx &= \int_{-5}^{0} 0.05x\,dx + \int_{0}^{5} 0.15x\,dx \\
&= \frac{0.05}{2}x^2\Big|_{-5}^{0} + \frac{0.15}{2}x^2\Big|_{0}^{5} \\
&= [0.025 \cdot 0 - 0.025 \cdot (-5)^2] + [0.075 \cdot 5^2 - 0.075 \cdot 0] \\
&= -0.625 + 1.875 \\
&= 1.25 = \frac{5}{4} < q_2 = \frac{5}{3}
\end{aligned}$$

[Use the property of expected values] From the pdf $f(x)$, there is a 0.25 chance that X is between [-5, 0) uniformly, and 0.75 chance that X is between [0, 5] uniformly. So X can be considered as a weighted average of two uniform random variables.

$$X = 0.25X_1 + 0.75X_2, \text{ where } X_1 \sim Unif[-5, 0), X_2 \sim Unif[0, 5]$$

Then

$$E(X) = 0.25E(X_1) + 0.75E(X_2) = 0.25 \cdot (-2.5) + 0.75 \cdot 2.5 = -0.625 + 1.875 = 1.25.$$

$$E(x) = 1.25 = \frac{5}{4} < q_2 = \frac{5}{3}$$