# Densities and Distributions

Jacek Wallusch

Instytut Kliometrii i Badań nad Transformacją

Probability density function is used to calculate the probability. Calculations require solving an integral of a following form:

$$\begin{aligned}\mathsf{Prob}\left(\alpha \leq X \leq \beta\right) &= \mathsf{Prob}\left(\beta \leq X\right) - \mathsf{Prob}\left(\alpha \leq X\right) = \\ &= \mathsf{F}\left(\beta\right) - \mathsf{F}\left(\alpha\right) = \\ &= \int_{\alpha}^{\beta} f\left(x\right) dx,\end{aligned}$$

where $f\left(x\right)$ is the probability density function.

# Problem

Suppose that the PDF of a random variable $x$ is:

$$f(x) = \begin{cases} x + \beta & \forall x \in [1, 2] \\ 0 & \text{elsewhere} \end{cases} \tag{1}$$

and we would like to calculate the probability $\text{Prob}(1 \leq x \leq 1.5)$. The solution requires:

1. To solve the integral
2. To specify the $\beta$
3. To substitute $\beta$ into the integral and to solve it for upper and lower limits equal to 1.5 and 1, respectively.

# Step 1: Solving the Integral

$$
\begin{aligned}
F(x) &= \int_{-\infty}^{\infty} x + \beta dx \\
&= \int_{1}^{2} x + \beta dx \\
&= \frac{x^2}{2} + \beta x \Big|_1^2 \\
&= \frac{2^2}{2} + 2\beta - \left(\frac{1^2}{2} + \beta\right) = \frac{4-1}{2} + \beta(2-1) \\
&= 1.5 + \beta.
\end{aligned}
$$

## Step 2: Finding the $\beta$

Set the solution to 1:

$$1.5 + \beta = 1, \tag{2}$$

and solve for $\beta$:

$$\beta = 1 - 1.5 = -0.5. \tag{3}$$

Question: Why should the solution be set to 1?

# Step 3: Finding the probability

$$\begin{aligned}\text{Prob}\,(1 \leq x \leq 1.5) &= \int_1^{1.5} x - 0.5dx \\ &= \frac{x^2}{2} - 0.5x\Big|_1^{1.5} \\ &= \frac{1.5^2}{2} - 0.5 \cdot 1.5 - \left(\frac{1^2}{2} - 0.5 \cdot 1\right) \\ &= 0.375 \qquad (4)\end{aligned}$$

Using the data (see the R-script http://wallusch-datenbank.de/resources/matstat/exp_dist.R) find the probability that the lifelength of a particular battery of this type is **less** than 200 or **greater** than 400 hours.

Schaeffer and McClave (1986), p. 60.

# Step 1: Histogram

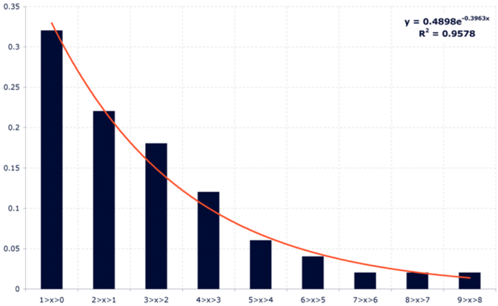
Frequency distribution

1. re-group the data in an ascending order
2. find any regularities
3. specify bins
4. perform histogram
5. calculate frequencies
6. graph the frequencies
7. assign a trend

# Step 2: Density function

Approximations

1. negative exponential curve
2. original function and its approximation $y = 0.5e^{-0.5x}$



What do we need the approximations for?
Calculate the definite integrals:

$$\int_0^9 0.5e^{-0.5x}dx$$

and

$$\int_0^9 0.4898e^{-0.3963x}dx$$

and then compare the results. Have you noticed anything important?
Improper integral: a limit of the following form

$$\lim_{B\to\infty}\int_A^B f(x)\,dx.$$

Using this property re-consider the approximation:

$$y=\int_0^\infty 0.5e^{-0.5x}dx$$

# Step 3: Re-Formulation of the Initial Problem

Using the density function

$$y = \begin{cases} 0.5e^{-0.5x} & \forall x > 0 \\ 0 & \text{elsewhere} \end{cases} \tag{5}$$

and noting that the events *less than 200* and *more than 400* are mutually exclusive, i.e.

$$\text{Prob}\,(A \cup B) = \text{Prob}\,(A) + \text{Prob}\,(B) \tag{6}$$

calculate the probability.

# Step 4: Solution

$$
\begin{aligned}
\text{Prob}\left(A \cup B\right) &= \text{Prob}\left(A\right) + \text{Prob}\left(B\right) \\
&= \int_0^2 0.5e^{-0.5x} dx + \int_4^\infty 0.5e^{-0.5x} dx \\
&= 1 - \exp\left(-1\right) + \exp\left(-2\right) \\
&= 0.767.
\end{aligned}
$$