Applied calculus *on-line chapter: calculus applied to probability and statistics*

Section 3. Exponential, Normal, and Beta Distributions

Note: The mathematics in this page has been typeset using *jsMath*. To see the *jsMath* typesetting at its best you should install the *jsMath* TeX fonts. Click on the *jsMath* button at the bottom right corner of this page for more details.

Exponential Density Function

You are an investment analyst, and you have information that mortgage lenders are failing continuously at 5% per year. What is the probability that a mortgage lender will fail sometime within the next x years?

To answer the question, suppose that you started with 100 mortgage lenders. Since they are failing continuously at a rate of 5% per year, the number left after x years is given by the decay equation

$$\text{Number left} = 100e^{-0.05x},$$

so

$$\begin{aligned}\text{Number that fail} &= \text{Total number} - \text{Number left}\\ &= 100 - 100e^{-0.05x}\\ &= 100(1 - e^{-0.05x}).\end{aligned}$$

Thus, the percentage that will have failed by that time—and hence the probability that we are asking for—is given by

$$P = \frac{100(1 - e^{-0.05x})}{100} = 1 - e^{-0.05x}.$$

Now let X be the number of years a randomly chosen mortgage lender will take to fail. We have just calculated the probability that X is between 0 and x. In other words,

$$P(0 \le X \le x) = 1 - e^{-0.05x}.$$

But we also know that

$$P(0 \le X \le x) = \int_0^x f(t)\,dt$$

for a suitable probability density function. Thus,

$$\int_0^x f(t)\,dt = 1 - e^{-0.05x}.$$

The Fundamental Theorem of Calculus tells us that the derivative of the left side is $f(x)$. Thus,

$$f(x) = \frac{d}{dx}\int_0^x f(t)\,dt = \frac{d}{dx}[1 - e^{-0.05x}] = 0.05e^{-0.05x},$$

which is the probability density function we were seeking.

Question Does this function satisfy the mathematical conditions necessary for it to be a probability density function?

Answer First, the domain of f is $[0, +\infty)$, since x refers to the number of years from now. Checking requirements (a) and (b) for a probability density function,

(a) $0.05e^{-0.05x} \ge 0$

$$\textbf{(b)} \int_0^{+\infty} 0.05e^{-0.05x}\,dx = \lim_{M \to +\infty} \int_0^M 0.05e^{-0.05x}\,dx$$
$$= \lim_{M \to +\infty} \left[-e^{-0.05x}\right]_0^M$$
$$= \lim_{M \to +\infty} (e^0 - e^{-0.05M}) = 1 - 0 = 1$$

There is nothing special about the number 0.05. Any function of the form

$$f(x) = ae^{-ax}$$

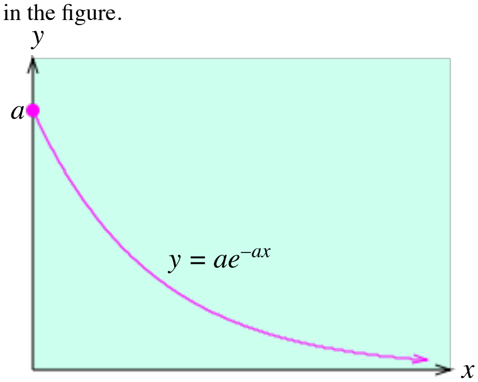
with a a positive constant is a probability density function. A density function of this form is referred to as an **exponential density function**.

Exponential Density Function

An **exponential density function** is a function of the form

$$f(x) = ae^{-ax} \qquad (a \text{ a positive constant})$$

with domain $[0 + \infty)$. Its graph is shown in the figure.



Top of Page

Example 1 *Failing Mortgage Lenders*

Continuing the example in the text, what is the probability that a given mortgage lender will fail between 2 and 4 years from now? What is the probability that it will last 5 or more years?

Solution Recall that our probability density function for the failure of mortgage lenders is given by

$$f(x) = ae^{-ax} = 0.05e^{-0.05x}$$

The probabilities we want are given by integrals.

$$P(2 \le X \le 4) = \int_2^4 0.05e^{-0.05x}\,dx$$
$$= \left[-e^{-0.05x}\right]_2^4$$
$$= e^{-0.01} - e^{-0.02} \approx .086.$$

$$P(X \ge 5) = \int_5^{+\infty} 0.05e^{-0.05x}\,dx$$
$$= \lim_{M \to +\infty} \int_5^M 0.05e^{-0.05x}\,dx$$
$$= \lim_{M \to +\infty} \left[-e^{-0.05x}\right]_5^M$$
$$= \lim_{M \to +\infty} (e^{-0.25} - e^{-0.05M})$$
$$= -e^{-0.25} \approx .779.$$

So there is an 8.6% chance that a given mortgage lender will fail between 2 and 4 years from now, and a 77.9% chance that it will last 5 or more years.

Before We Go On ... We could also calculate $P(X \geq 5)$ as $1 - P(0 \leq X \leq 5)$ and thus avoid having to calculate an improper integral. an imporper integral. Try this for practice.

Top of Page

Example 2 *Radioactive Decay*

Plutonium 239 decays continuously at a rate of 0.00284% per year. If X is the time a randomly chosen plutonium atom will decay, write down the associated probability density function, and use it to compute the probability that a plutonium atom will decay between 100 and 500 years from now.

Solution Using the discussion on failing mortgage lenders as our guide, we see that $a = 0.0000284$, so that the probability density function is

$$f(x) = ae^{-ax} = 0.0000284e^{-0.0000284x}$$

For the second part of the question,

$$P(100 \leq X \leq 500) = \int_{100}^{500} 0.0000284e^{-0.0000284x}\, dx \approx .011.$$

Thus, there is a 1.1% chance that a plutonium atom will decay sometime during the given 400 year period.

Top of Page

Normal Density Function

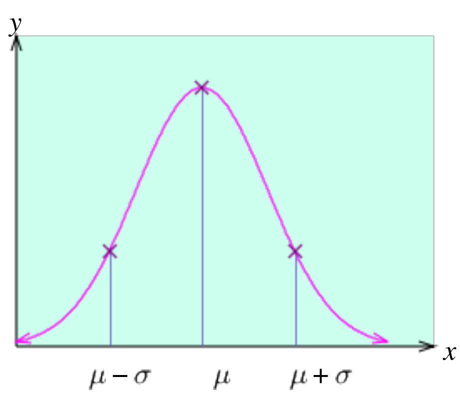
Perhaps the most interesting class of probability density functions are the **normal density functions**, defined as follows.

Normal Density Function

A **normal density function** is a function of the form

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}},$$

with domain $(-\infty, +\infty)$. The quantity μ is called the **mean** and can be any real number, while σ is called the **standard deviation** and can be any positive real number. The graph of a normal density function is shown in the following figure:



Properties of the Normal Density Curve

You can check the following properties using calculus and a little algebra:

(1) The normal curve is "bell-shaped" with the maximum occurring at $x = \mu$ (center point marked on the graph).

(2) It is symmetric about the vertical line $x = \mu$.

(3) It is concave down in the range $\mu - \sigma \leq x \leq \mu + \sigma$, and concave up outside that range.

(4) There are inflection points at $x = \mu - \sigma$ and $x = \mu - \sigma$ as marked on the graph.

(5) (Less obvious) The integral of the normal density function is given in terms of the Gauss error function:

$$\int f(x)\,dx = \frac{1}{2}\mathrm{erf}\left(\frac{x-\mu}{\sqrt{2}\sigma}\right) + C$$

The normal density function applies in many situations that involve measurement and testing. For instance, repeated imprecise measurements of the length of a single object, a measurement made on many items from an assembly line, and collections of SAT scores tend to be distributed normally. It is for this reason that the normal density curve is so important in quality control and in assessing the results of standardized tests.

In order to use the normal density function to compute probabilities, we need to calculate integrals of the form $\int_a^b f(x)\,dx$. However, as we saw above, the antiderivative of the normal density function cannot be expressed in terms of any commonly used functions. Traditionally, statisticians and others have used tables coupled with transformation techniques to evaluate such integrals. This approach is rapidly becoming obsolete as the technology of hand-held computers and programmable calculators puts the ability to do numerical integration quickly and accurately in everybody's hands (literally). In keeping with this trend, we shall show how to use various technologies do the necessary calculation in the next example.

Top of Page

Example 3 *Quality Control*

Pressure gauges manufactured by Precision Corp. must be checked for accuracy before being placed on the market. To test a pressure gauge, a worker uses it to measure the pressure of a sample of compressed air known to be at a pressure of exactly 50 pounds per square inch. If the gauge reading is off by more than 1% (0.5 pounds), the guage is rejected. Assuming that the reading of a pressure gauge under these circumstances is a normal random variable with mean 50 and standard deviation 0.5, find the percentage of gauges rejected.

Solution For a gauge to be accepted, its reading X must be 50 to within 1%, in other words, $49.5 \le X \le 50.5$. Thus, the probability that a gauge will be accepted is $P(49.5 \le X \le 50.5)$. X is a normal random variable with $\mu = 50$ and $\sigma = 0.5$. The formula tells us that

$$P(49.5 \le X \le 50.5) = \int_{49.5}^{50.5} f(x)\,dx.$$

where

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(x-\mu)^2}{2\sigma^2}} = \frac{1}{0.5\sqrt{2\pi}}e^{-\frac{(x-50)^2}{2(0.5)^2}},$$

Method 1: Calculating the intergal numerically Using a TI-83/4, for example, you would enter the above formula as

```
Y1=(1/(0.5(2π)^0.5))e^(-(X-50)^2/0.5)
```

and then enter

```
fnInt(Y1,X,49.5,50.5)
```

in the home screen. Alternatively, the built-in `normalcdf` function in the TI-83/4 permits one compute $P(a \le X \le b)$ directly as well. The format for this is

```
normalcdf(a, b, μ, σ)
```

In Excel, $P(a \le X \le b)$ is computed with the formula

```
=NORMDIST(b, μ, σ 1)-NORMDIST(a, μ, σ 1)
```

Alternatively, you could use the Numerical integration utility on this Website: Enter the formula

```
Y1=(1/(0.5(2*pi;)^0.5))e^(-(X-50)^2/0.5)
```

for "$f(x)$," adjust the accuracy as desired, and press "Adaptive Quadrature."

The numerical calculation yields an answer of approximately 0.6827. In other words, 68.27% of the gauges will be accepted. Thus, the remaining 31.73% of the gauges will be rejected.

Before We Go On ... Here is a utility that calculates the area under the normal curve with high accuracy (to around 16 decimal places). The algorithm used to compute the probabilities is from a collection of powerful statistical algorithms due to Ian Smith:

Notes:

1. You must fill in values for the mean and standard deviation.

2. To compute, say, $P(X \ge 1.2)$ enter $P(1.2 \le X \le \quad)$

3. To compute, say, $P(X \le 1.2)$ enter $P(\quad \le X \le 1.2)$

Mean μ = 0	St. Deviation σ = 1	Rounding: 6
$P($ 0 $\le X \le$ 1 $)$	Calculate Probability	Table Please
Clear		

Normal Probability Calculator

As we mentioned above, the traditional and still common way of calculating normal probabilities is to use tables. The tables most commonly published are for the **standard normal distribution**, the one with mean 0 and standard deviation 1. If X is a normal variable with mean μ and standard deviation σ, the variable $Z = (X - \mu)/\sigma$ is a standard normal variable (see the exercises). Thus, to use a table we first write

$$P(a \le X \le b) = P\left[\frac{a-\mu}{\sigma} \le Z \le \frac{b-\mu}{\sigma}\right]$$

and then use the table to calculate the latter probability (Z always stands for the standard normal variable).

The following calculations, true for any normal random variable, are very useful to remember:

$$P(\mu - \sigma \le X \le \mu + \sigma) \approx .6827$$
$$P(\mu - 2\sigma \le X \le \mu + 2\sigma) \approx .9545$$
$$P(\mu - 3\sigma \le X \le \mu + 3\sigma) \approx .99737$$

Question Why can we assume that the reading of a pressure gauge is given by a normal distribution? Why is the normal distribution so common in this kind of situation?

Answer The reason for this is rather deep. There is a theorem in probability theory called the Central Limit Theorem that says that a large class of probability density functions may be approximated by normal density functions. Repeated measurement of the same quantity gives rise to such a function.

Top of Page

Beta Density Function

There are many random variables whose values are percentages or fractions. These variables have density functions defined on $[0, 1]$. A large class of random variables, such as the percentage of new businesses that turn a profit in their first year, the percentage of banks that default in a given year, and the percentage of time a plant's machinery is inactive, can be modeled by a **beta density function**.

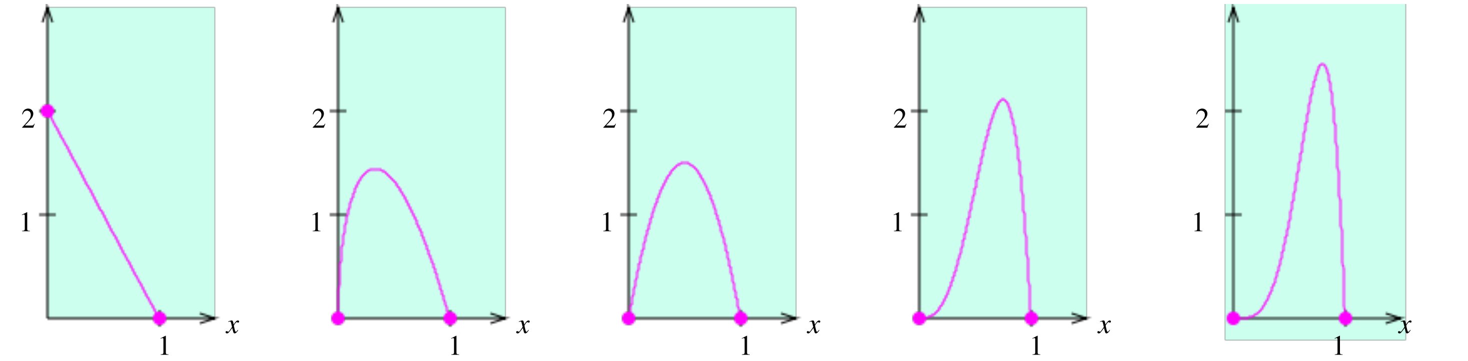
Beta Density Function

A **beta density function** is a function of the form

$$f(x) = (\beta + 1)(\beta + 2)x^{\beta}(1 - x)$$

with domain $[0, 1]$. The number β can be any nonnegative constant. Below are the graphs of $f(x)$ for several values of β. You can adjust the value of β in the last one (change the value and press "Return" or "Enter").

y $\quad$ y $\quad$ y $\quad$ y $\quad$ y



$\beta = 0$	$\beta = 0.5$	$\beta = 1$	$\beta = 3$	$\beta =$ 4 Clear
$f(x) = 2(1-x)$	$f(x) = 3.75x^{0.5}(1-x)$	$f(x) = 6x(1-x)$	$f(x) = 20x^3(1-x)$	$f(x) = (\beta+1)(\beta+2)x^{\beta}(1-x)$

Example 4 *Downsizing in the Utilities Industry*

A utilities industry consultant predicts a cutback in the Canadian utilities industry during 2010-2015 by a percentage specified by a beta distribution with $\beta = 0.25$. Calculate the probability that Ontario Hydro will downsize by between 10% and 30% during the given five-year period.

Solution The beta density function with $\beta = 0.25$ is

$$\begin{aligned} f(x) &= (\beta+1)(\beta+2)x^{\beta}(1-x) \\ &= 2.8125x^{0.25}(1-x) \\ &= 2.8125(x^{0.25} - x^{1.25}). \end{aligned}$$

Thus,

$$\begin{aligned} P(0.10 \le X \le 0.30) &= \int_{0.10}^{0.30} 2.8125(x^{0.25} - x^{1.25})\, dx \\ &= 2.8125\left[\frac{x^{1.25}}{1.25} - \frac{x^{2.25}}{2.25}\right]_{0.10}^{0.30} \approx 0.2968. \end{aligned}$$

So there is approximately a 30% chance that Ontario Hydro will downsize by between 10% and 30%.

Before We Go On ... Go to the definition box above and put $\beta = 0.25$ to see the graph of the associated density function. You will notice that its shape is "in between" those for $\beta = 0$ and $\beta = 0.5$ shown on the left.

Top of Page