LECTURE: NORMAL DISTRIBUTION

1. Normal Distribution (continued)

Recall:

The **normal distribution** is $Y \sim \text{N}(\mu, \sigma)$. The pdf of Y is

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

Recall:

The **standard normal distribution** is $Z \sim N(0, 1)$ with pdf

$$f(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}$$

$Z-$table:

Let $Z \sim N(0, 1)$ and suppose we wish to calculate $P(-1 \leq Z \leq 1)$. Using the density of the standard normal:

$$P(-1 \leq Z \leq 1) = \int_{-1}^{1} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz$$

Unfortunately, there is no nice antiderivative for the integrand, so we cannot compute the integral using the fundamental theorem of calculus. One option is to use numerical integration techniques, which can be done using software packages such as Matlab or Mathematica.

Another option is to use tables for the cdf of the standard normal distribution. Although this is perhaps a bit "old-school," it is important you know how to use these tables.

There are many versions of the Z-table. The one that will be provided (see course website) is the actual cdf for Z, i.e. it gives values for $F(z) = P(Z \leq z)$. Since the standard normal is symmetric about 0, some tables only provide values on one side of the mean, since the others can be computed using symmetry. The table in the textbook by Wackerly et al, for example, provides $P(Z \geq z)$ for $z \geq 0$.

How do we compute this using a Z-table? Letting $F(z)$ be the cdf for the standard normal distribution, recall that for any continuous probability distribution we have:

$$P(a \leq Z \leq b) = F(b) - F(a)$$

Here $P(-1 \leq Z \leq 1) = F(1) - F(-1) = 0.8413 - 0.1587 = 0.6826$

Example 1:

Let $Z \sim N(0,1)$ and let $F(z)$ be its cdf.

(a) Find $P(Z > 2)$

$$P(Z > 2) = 1 - P(Z \leq 2) = 1 - F(2) = 1 - 0.9772 = 0.0228$$

(b) Find $P(-2 \leq Z \leq 2)$

$$P(-2 \leq Z \leq 2) = F(2) - F(-2) = 0.9772 - 0.0228 = 0.9544$$

(c) Find $P(0 \leq Z \leq 1.73)$

$$P(0 \leq Z \leq 1.73) = F(1.73) - F(0) = 0.9582 - 0.5 = 0.4582$$

Here we use $F(0) = P(Z \leq 0) = 0.5$ which is true because the normal distribution is symmetric about 0

68-95-99 Rule:

Since the standard deviation of Z is 1, then $P(-1 \leq Z \leq 1) \approx 0.68$ is the probability of falling within one standard deviation of the mean and $P(-2 \leq Z \leq 2) \approx 0.95$ is the probability of falling within to standard deviations of the mean. These are useful numbers to remember, since they are good guidelines for interpreting the normal distribution.

68-95-99 Rule:

Let Y be a random variable with a normal distribution. Then:

(1) The probability of falling within 1 standard deviation of the mean is about 0.68

(2) The probability of falling within 2 standard deviations of the mean is about 0.95

(3) The probability of falling within 3 standard deviations of the mean is about 0.997

Shifting:

What if we don't have a standard normal random variable?

Fun Fact:

We can transform *any* normal variable $Y \sim \mathrm{N}(\mu, \sigma)$ into a standard normal random variable $Z \sim \mathrm{N}(0, 1)$ using

$$Z = \frac{Y - \mu}{\sigma}$$

In other words, subtract the mean and divide by the standard dev.

Example 2:

A machine produces balls which diameters that are normally distributed with mean 3.0005 cm and standard deviation 0.0010 cm. Specifications require the diameters to lie in the interval 3.0000 ± 0.0020 cm. What fraction of the total production meets those specifications?

Let Y be the diameter of a ball, then $Y \sim \mathrm{N}(3.0005, 0.0010)$

Here we want $P(2.9980 \leq Y \leq 3.0020)$

To use the $z-$table, first convert $Y \sim \mathrm{N}(\mu, \sigma)$ to $Z \sim \mathrm{N}(0, 1)$:

$$y = 2.9980 \qquad z = \frac{2.9980 - 3.0005}{0.0010} = -2.5$$

$$y = 3.0020 \qquad z = \frac{3.0020 - 3.0005}{0.0010} = 1.5$$

$$\begin{aligned} P(2.9980 \leq Y \leq 3.0020) =& P(-2.5 \leq Z \leq 1.5) = F(1.5) - F(-2.5) \\ =& 0.9332 - 0.0062 = 0.9270 \end{aligned}$$

So approximately 92.7% of the balls meet the required specifications.

2. Exponential Distribution

The final continuous distribution we will discuss is the exponential distribution. It belongs to the family of gamma distributions, and is perhaps the most useful member of that family, it is the only one we'll consider in this class.

The exponential distribution is used to model the length of time between independent event with constant average rate. Think for instance the length of time between customer arrivals at a restaurant or phone calls at a call center.

Notice those are the same events as for the Poisson distribution!

(1) The Poisson distribution (discrete) measures the number of events which occur in a fixed span of time.

(2) The exponential distribution (continuous) measures the amount of time between two subsequent events.

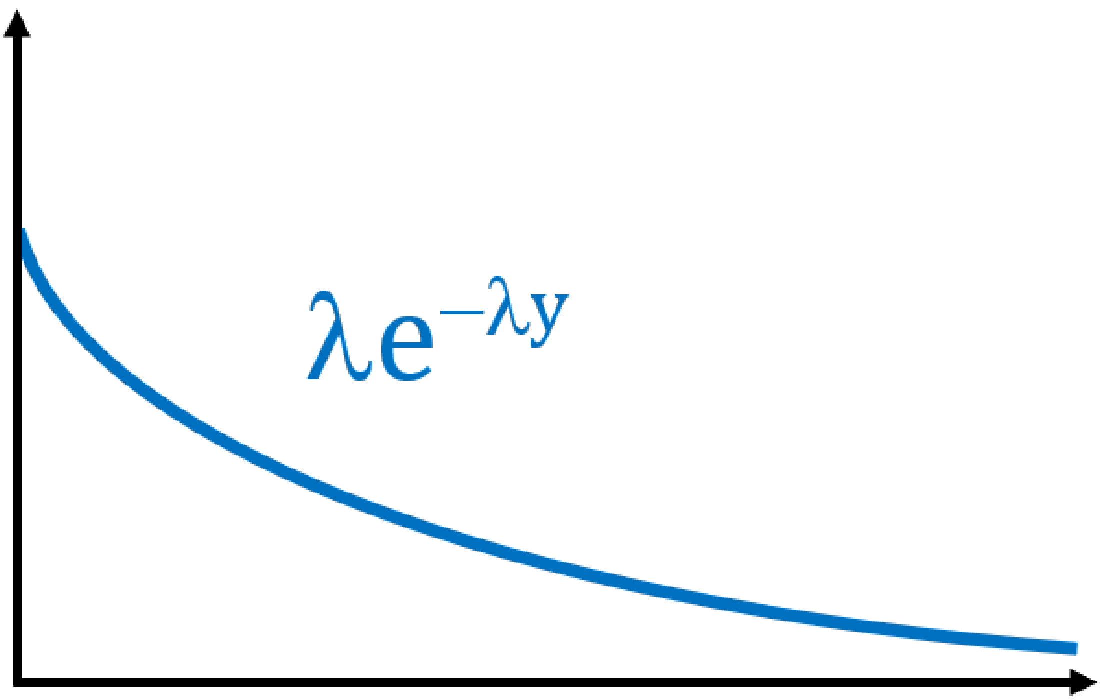
It also used to model the lifetime of electronic and mechanical components, similar to the computer crash problem in the geometric distribution.

Definition:

Y has an **exponential distribution** with parameter $\lambda > 0$ if

$$f(y) = \begin{cases} \lambda e^{-\lambda y} & y \geq 0 \\ 0 & y < 0 \end{cases}$$

We write $Y \sim \text{Exp}(\lambda)$



As always, we verify that the exponential distribution integrates to 1

$$\int_{-\infty}^{\infty} f(y)dy = \int_0^{\infty} \lambda e^{-\lambda y} dy \overset{\text{DEF}}{=} \lim_{t\to\infty} \int_0^t \lambda e^{-\lambda y} dy = \lambda \lim_{t\to\infty} \left[\frac{\lambda e^{-\lambda y}}{-\lambda}\right]_0^t$$
$$= -\left(\lim_{t\to\infty} e^{-\lambda t} - 1\right) = -(0-1) = 1$$

Fact:

Suppose $Y \sim$ Exp (λ) then $E(Y) = \frac{1}{\lambda}$ and $\text{Var}(Y) = \frac{1}{\lambda^2}$

$$\begin{aligned}
E(Y) &= \int_{-\infty}^{\infty} y f(y) dy \\
&= \int_0^{\infty} \lambda y e^{-\lambda y} dy \\
&= \lambda \lim_{t\to\infty} \int_0^t y e^{-\lambda y} dy \\
&\overset{\text{IBP}}{=} \lambda \lim_{t\to\infty} \left(\left[-\frac{1}{\lambda} y e^{-\lambda y} \right]_0^t + \frac{1}{\lambda} \int_0^t e^{-\lambda y} dy \right) \\
&= -\lim_{t\to\infty} t e^{-\lambda t} + 0 + \frac{1}{\lambda} \underbrace{\int_0^{\infty} e^{-\lambda y} dy}_{1} \\
&= \frac{1}{\lambda}
\end{aligned}$$

Where in the second-to-last line the limit is 0 by L'Hôpital's rule.

Similarly, using the Magic Variance Formula and integrating by parts twice, we can prove the variance of an exponential random variable.