STAT 509 – Section 3.4: Continuous Distributions

Probability distributions are used a bit differently for continuous r.v.'s than for discrete r.v.'s.

- **A continuous random variable is one for which the outcome can be any value in an interval of the real number line.**
- **Examples**
  - **Let Y = length in mm**
  - **Let Y = time in seconds**
  - **Let Y = temperature in ºC**

Continuous distributions typically are represented by a <u>probability density function</u> (pdf), denoted $f(y)$.

<u>Properties of Density Functions:</u>

(1) $f(y) \geq 0$ for all possible y values. (Density function always on or above the horizontal axis)

(2) $\int_{-\infty}^{\infty} f(y)dy$ = 1. (Total area beneath the curve is exactly 1.)

(3) The cumulative distribution function (cdf) is again denoted by F(y).

$F(y) = P(Y \leq y) = \int_{-\infty}^{y} f(t)dt$

(4) For continuous r.v.'s, the probability distribution will give us the probability that a value falls in an <u>interval</u> (for example, between two numbers).

That is, the probability distribution of a continuous r.v. Y will tell us $P(y_1 \le Y \le y_2)$, where y_1 and y_2 are particular numbers of interest.

Specifically, $P(y_1 \le Y \le y_2)$ is the *area under the density function* between $y = y_1$ and $y = y_2$:

$$P(y_1 \le Y \le y_2) = \int_{y_1}^{y_2} f(y)\,dy = F(y_2) - F(y_1)$$

Example Picture:

Example 1: Suppose *Y*, the grams of lead per liter of gasoline, has the density function:

$$f(y) = 12.5y - 1.25 \qquad \text{for } 0.1 \leq y \leq 0.5$$

What is the probability that the next liter of gasoline has less than 0.3 grams of lead?

Find the cdf for this r.v. *Y*.

In general, the expected value of a continuous r.v. *Y* is

μ = E(*Y*) = $\int_{-\infty}^{\infty} yf(y)dy$

and the variance of a continuous r.v. *Y* is

σ^2 = $\int_{-\infty}^{\infty} (y-\mu)^2 f(y)dy$

Find the expected value of the r.v. in Example 1.

Nice exercise: Find the variance of the r.v. in Example 1.

The Uniform Distribution

This is a simple example of a continuous distribution.

A uniform r.v. is equally likely to take any value between its lower limit (some number a) and its upper limit (some number b).

Density looks like a rectangle:

If total area is 1, then what is the height of the density function?

- **Mean of a Uniform(a, b) r.v. = $(a + b) / 2$**
- **Variance of a Uniform(a, b) r.v. = $(b - a)^2 / 12$**

Example: A machine designed to fill 16-ounce water bottles actually dispenses a random amount between 15.0 and 17.0 ounces. The amount Y of water dispensed is a Uniform(15, 17) random variable:
Density:

What is the probability that the bottle has less than 15.5 ounces of water?

P(Y < 15.5) = P(15 < Y < 15.5) =

The Exponential Distribution

Often, waiting times can be modeled with an exponential distribution:

pdf: $f(y) = \lambda e^{-\lambda y}$, **for $y > 0$ and $\lambda > 0$**

• **Here, λ is the rate of the exponential distribution.**

For example, suppose Y is the waiting time (in days) between accidents at a plant. Then λ is the expected number of accidents per day.

The mean waiting time is

μ = E(Y) = $\int_0^\infty y f(y)\,dy$ **= 1 / λ (Proof?)**

The variance is 1 / λ^2 → Standard deviation =

Example: Suppose a plant has a rate of 0.05 accidents per day and that the waiting time between accidents follows an exponential distribution.

What is the probability that the *time between* the next two accidents will be less than 10?

Picture:

Solution: P(*Y* < 10) = F(10) =

In general: F(*y*) = P(*Y* ≤ *y*) = $\int_0^y \lambda e^{-\lambda t}\,dt$ **=**

And so P(*Y* > *y*) =

Picture again:

Example: If the time to failure for an electrical component follows an exponential distribution with a mean failure time of 1000 hours. What is the probability that a randomly chosen component will fail before 750 hours?

Relationship between Exponential and Poisson r.v.'s

- **Suppose the Poisson distribution is used to model the probability of a specific number of events occurring in a particular interval of time or space.**
- **Then the time or space between events is an *exponential* random variable.**

Why? Suppose we have a Poisson distribution with mean λt, where t is a particular length of time.

What is the probability that no events will occur before time t?

Note: This situation implies that the waiting time T before the next event is <u>more than</u> t.

Therefore $P[T > t] =$

And so T must follow an exponential distribution with rate λ.

Example: Suppose the number of machine failures in a given interval of time follows a Poisson distribution with an average of 1 failure per 1000 hours.

What is the probability that there will be no failures during the next 2000 hours?

What is the probability that the time until the next failure is more than 2000 hours?

What is the probability that more than 2500 hours will pass before the next failure occurs?

• The exponential distribution is sometimes not flexible enough to model realistic waiting time distributions.

• The Weibull distribution is a more flexible, two-parameter waiting-time distribution with pdf:

$f(y) = \lambda\beta(\lambda y)^{\beta-1} e^{-(\lambda y)^{\beta}}$, **for $y > 0$ and $\lambda, \beta > 0$**

• Note the exponential distribution is simply the Weibull with β =

• The gamma distribution is another two-parameter generalization of the exponential distribution (see Table 3.5 for pdf, mean, variance)