4 Continuous Random Variables

Chapter Preview

In this chapter, we'll introduce continuous random variables which are random variables that take on possible values from an uncountable set (e.g., interval or a union of disjoint intervals). We will also look at some common continuous random variables such as the Normal distribution.

4.1: Probability Density Functions (pdf)

Recall: A continuous random variable is a random variable whose possible values either constitute an interval of real numbers or a union of intervals of real numbers.

Def: Let X be a continuous random variable. The probability distribution or probability density function (pdf) of X is a function $f(x)$ such that for any two numbers a and b with $a \leq b$, $P(a \leq X \leq b) = \int_a^b f(x)dx$.

Note: All valid pdfs must satisfy two conditions:
(1) $f(x) \geq 0$ ($f(x)$ is non-negative).
(2) $\int_{-\infty}^{\infty} f(x)dx = 1$ (area under f(x) is equal to 1).

Example: Let X be the IQ of a randomly chosen individual which follow a normal distribution with mean 100 and standard deviation 15.

Example: Let Y be a random real number between 0 and 1. (Note: this can be done *approximately* in R using the **runif** function.)

Example: Suppose X is a random variable with the following pdf:

$$f(x) = \begin{cases} 0 & x < 0 \\ \frac{x}{2} & 0 \leq x \leq 2 \\ 0 & x > 2 \end{cases}$$

(a) Is $f(x)$ a valid pdf?

(b) What is $P(X < 1)$?

(c) What is $P(1/2 \leq X \leq 3/2)$?

(d) What is $P(X = 1)$?

Caution! For a continuous random variable X, $P(X = c) = 0$ for all constants c. While it may be possible for X to be exactly equal to the constant c, it can only happen with probability zero. In general, when dealing with continuous random variables, we consider the probability that X takes on a value over an interval rather than a specific value.

4.2: Cumulative Distribution Functions and Expected Values

Cumulative Distribution Functions

Def: The cumulative distribution function (CDF) of a continuous random variable X is defined for every number x by $F(x) = P(X \leq x) = \int_{-\infty}^{x} f(y)dy$. F(x) represents the area under the curve to the left of x.

Note: Conditions on $F(x)$:
(1) $0 \leq F(x) \leq 1$
(2) $\lim_{x \to -\infty} F(x) = 0$
(3) $\lim_{x \to \infty} F(x) = 1$

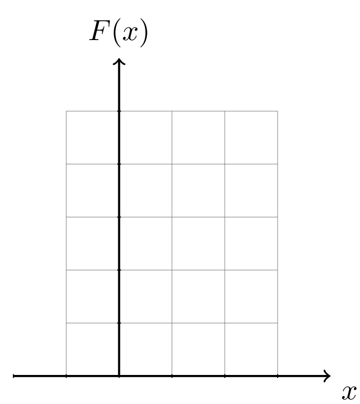
Theorem: If X is a continuous random variable with pdf $f(x)$ and CDF $F(x)$, then at every x at which the derivative $F'(x)$ exists, $F'(x) = f(x)$.

Theorem: Let X be a continuous random variable with pdf $f(x)$ and CDF $F(x)$. Then we have the following:
(1) For any number a, $P(X > a) = 1 - F(a)$
(2) For any two numbers a and b with $a \leq b$, $P(a \leq X \leq b) = F(b) - F(a)$.

Example: Suppose X is a random variable with the following pdf:

$$f(x) = \begin{cases} 0 & x < 0 \\ \frac{x}{2} & 0 \leq x \leq 2 \\ 0 & x > 2 \end{cases}$$

Find the CDF of X and sketch it. Also, calculate $P(X \leq 1)$ using the CDF.



Expected Values

Def: The expected value or mean of a continuous random variable X with pdf $f(x)$ is

$$E[X] = \mu = \int_{-\infty}^{\infty} x \cdot f(x)dx$$

Theorem: If X is a continuous random variable with pdf $f(x)$ and $h(X)$ is any function of X, then

$$E[h(X)] = \int_{-\infty}^{\infty} h(x) \cdot f(x)dx$$

Def: The variance of a continuous random variable X with pdf $f(x)$ is

$$Var(X) = \sigma^2 = \int_{-\infty}^{\infty} (x-\mu)^2 \cdot f(x)dx = E[(X-\mu)^2] = E[X^2] - (E[X])^2$$

Def: The standard deviation of a continuous random variable X is $\sigma = \sqrt{Var(X)}$.

Theorem: If X is a continuous random variable and $Y = aX + b$, then $E[Y] = aE[X] + b$ and $Var(Y) = a^2 Var(X)$.

Example: Suppose X is a random variable with the following pdf:

$$f(x) = \begin{cases} 0 & x < 0 \\ \frac{x}{2} & 0 \leq x \leq 2 \\ 0 & x > 2 \end{cases}$$

(a) Calculate the expected value of X.

(b) Calculate the expected value of X^2.

(c) Calculate the variance of X.

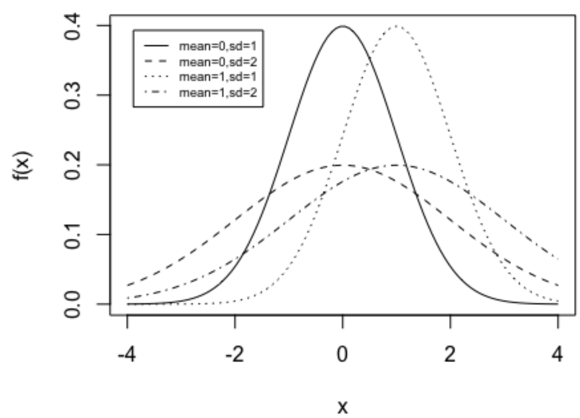
4.3: The Normal Distribution

Def: A continuous random variable X is said to be Normal(μ,σ^2) or $\mathcal{N}(\mu, \sigma^2)$ random variable if the pdf of X is:

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

.

If $X \sim \mathcal{N}(\mu, \sigma^2)$, then $E[X] = \mu$ and $Var(X) = \sigma^2$.



Def: The normal distribution with parameter values $\mu = 0$ and $\sigma = 1$ is called the standard normal distribution.

Note: $X \sim \mathcal{N}(\mu, \sigma^2)$, F(x) is denoted $\Phi(x)$. There is no closed form solution for $\Phi(x)$ since $\Phi(x) = F(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(x-\mu)^2/(2\sigma^2)} = (?)$. Instead, we use a table called a standard normal table to assist in these calculations.

Theorem: If $X \sim \mathcal{N}(\mu, \sigma^2)$, that is, X is a normal distribution with mean μ and standard deviation σ, then $Z = \frac{X-\mu}{\sigma}$ has a standard normal distribution.

Note: This normalization method allows us to calculate probability for any normal distribution using the standard normal table.

Example: Suppose $Z \sim \mathcal{N}(0, 1)$, that is, Z is a standard normal distribution. Calculate the following and draw an accompanying picture:

(a) $P(Z < 0.68)$

(b) $P(Z \geq -1.21)$

(c) $P(0.43 \leq Z < 0.68)$

(d) Suppose $P(Z < c) = 0.89$. Find c.

Example: Suppose the weights of house cats are approximately normally distributed with mean of 10 pounds and a standard deviation of 3 pounds. Calculate the following and draw an accompanying picture:

(a) $P(X < 15)$

(b) What is the probability that a randomly chosen house cat will weigh between 8 and 12 pounds?

(c) What is the 99^{th} percentile of house cat weights?

Theorem: The Empirical Rule states that if X is (approximately) normally distributed then:
(1) Approximately 68% of the values are within 1 SD of the mean.
(2) Approximately 95% of the values are within 2 SD of the mean.
(3) Approximately 99.7% of the values are within 3 SD of the mean.

Example: Suppose the weights of house cats are approximately normally distributed with mean of 10 pounds and a standard deviation of 3 pounds.

(a) Between what two bounds do approximately 68% of house cat weights lie?

(b) Between what two bounds do approximately 95% of house cat weights lie?

(a) Between what two bounds do approximately 99.7% of house cat weights lie?

4.4: The Exponential and Uniform Random Variables

Def: A continuous random variable X is said to be an Exponential(λ) random variable or an exponential random variable with parameter $\lambda > 0$ if the pdf of X is:

$$f(x|\lambda) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \\ 0 & otherwise \end{cases}$$

If $X \sim exp(\lambda)$, then $E[X] = \frac{1}{\lambda}$ and $Var(X) = \frac{1}{\lambda^2}$.

Example: Suppose $X \sim exp(\lambda)$, find the CDF of X.

Example: Show that if $X \sim exp(\lambda)$, then $E[X] = \frac{1}{\lambda}$. (Note: this requires integration by parts).

(Continuous) Uniform[a,b] Random Variable

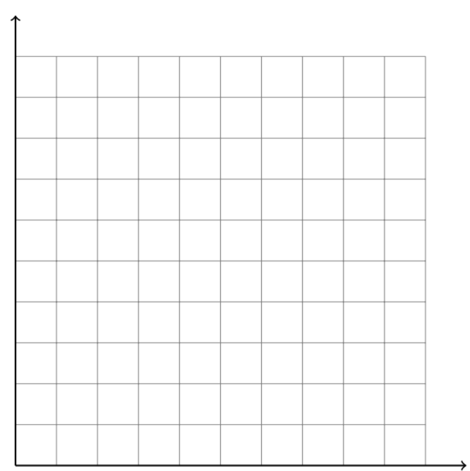
<u>Def:</u> A continuous random variable X is said to be <u>Uniform[a,b]</u> random variable if the pdf of X is:

$$f(x|a,b) = \begin{cases} \frac{1}{b-a} & a \leq x \leq b \\ 0 & otherwise \end{cases}$$

If $X \sim$ Uniform(a,b) random variable, then $E[X] = \frac{a+b}{2}$ and $Var(X) = \frac{(b-a)^2}{12}$.

<u>Example:</u> Show that if $X \sim Uniform[a,b]$, then $E[X] = \frac{a+b}{2}$.

<u>Example:</u> Suppose your local bus shows up once every 60 minutes, but you don't have a schedule, so you assume it's equally likely to show up at any time in the next 60 minutes. Let X be the amount of time you have to wait for the bus. Sketch the pmf and calculate the probability that your bus shows up in the next 10 minutes.



Common Continuous Random Variables summary table

Random Variable	**PMF**	$\mathbf{E}[\mathbf{X}]$	$\mathbf{Var}(\mathbf{X})$
Exponential(λ)	$f(x) = \lambda e^{-\lambda x}, 0 \leq x < \infty$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$
Normal(μ, σ^2)	$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(x-\mu)^2/(2\sigma^2)}, -\infty < x < \infty$	μ	σ^2
Uniform(a, b)	$f(x) = \frac{1}{b-a}, a \leq x \leq b$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$