Continuous Probability Distributions

Continuous Random Variables and Probability Distributions

- Random Variable: Y
- Cumulative Distribution Function (CDF): $F(y)=P(Y\leq y)$
- Probability Density Function (pdf): $f(y)=dF(y)/dy$
- Rules governing continuous distributions:
  - $f(y) \geq 0 \;\forall\, y$
  - $\int_{-\infty}^{\infty} f(y)dy = 1$
  - $P(a\leq Y\leq b) = F(b)-F(a) = \int_a^b f(y)dy$
  - $P(Y=a) = 0 \;\forall\, a$

127

Expected Values

128

Expected Values

The expected value of a continupus random variable:

$$\mu_X = E(X) = \int_{-\infty}^{\infty} x \cdot f(x)\, dx.$$

If X is a continuous rv with pdf $f(x)$ and $h(X)$ is any function of X, then

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

For $h(X)$ a linear function, $E(aX+b) = aE(X)+b$.

129

Expected Values

The *variance* of a continuous random variable X with pdf $f(X)$ and mean value μ is

$$\sigma_X^2 = V(X) = E[(X-\mu)^2] = \int_{-\infty}^{\infty} (x-\mu)^2 f(x)\,dx$$

The *standard deviation* (SD) of X is $\sigma_X = \sqrt{V(X)}$. The variance or standard deviation tell us how "spread out" the distribution is.

$$V(X) = E\left(X^2\right) - [E(X)]^2$$

130

Expected Values

- Given a collection of n random variables $X_1, X_2, \ldots, X_n$ and n numerical constants $a_1, a_2, \ldots, a_n$, the random variable

$$Y = a_1X_1 + a_2X_2 + \cdots + a_nX_n$$

is called a linear combination of the X_is.

- Whether or not the X_is are independent,

$$E[a_1X_1 + a_2X_2 + \cdots + a_nX_n] = a_1E[X_1] + a_2E[X_2] + \cdots + a_nE[X_n]$$

131

Expected Values

- If $X_1, X_2, \ldots, X_n$ are independent with variances $\sigma_1^2, \sigma_2^2, \ldots, \sigma_n^2$ then

$$\begin{aligned} V(a_1X_1 + a_xX_2 + \cdots + a_nX_n) &= a_1^2V(X_1) + a_2^2V(X_2) + \cdots + a_n^2V(X_n) \\ &= a_1^2\sigma_1^2 + a_2^2\sigma_2^2 + \cdots + a_n^2\sigma_n^2 \end{aligned}$$

132

Uniform Random Variables: pdf

- We want to define a random variable X that is "equally likely" to take on any value in some finite interval (a,b). Formally this is nonsensical since the probability of a continuous random variable assuming a particular value is always 0. A better way of formalizing our intuition is that the probability of X falling in a subinterval of (a,b) should depend only on the length of the subinterval, not on its location within (a,b).

133

Uniform Random Variables: pdf

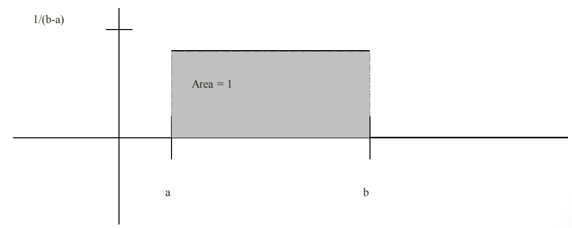
- The random variable X that satisfies this condition is the uniform random variable. We write X~uniform(a,b). It has the probability density function

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

134

Uniform Random Variables: pdf

- The graph of the pdf satisfies our intuition about "equal likelihood" of all intervals of a given length within (a,b). It also clearly has total area 1 under the pdf curve.



135

Uniform Random Variables: cdf

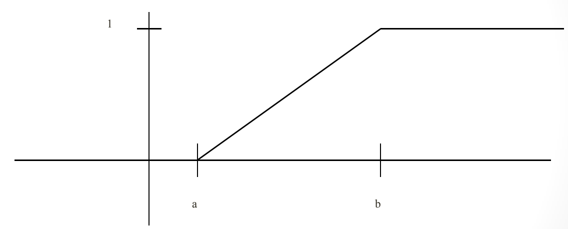
- The cumulative distribution function F given below is easy to compute either by integration of the pdf or by finding the area of rectangles. Note that it has all the usual properties of a cdf: 0 to the left, 1 to the right, increasing and continuous inbetween.

$$F(x) = \begin{cases} 0, & \text{if } x \le a \\ \frac{x-a}{b-a}, & \text{if } a < x < b \\ 1, & \text{if } x \ge b \end{cases}$$

136

Uniform Random Variables: cdf

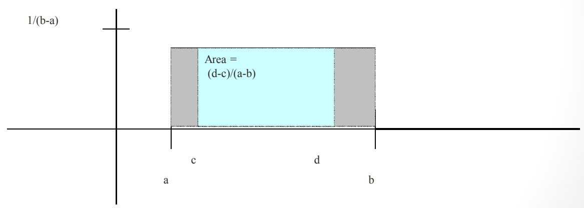
- Here we see F and its properties graphically.



137

Uniform Random Variables: cdf

- If we want to find the probability P(c<X<d) where a<c<d<b, then we can integrate formally, but it is easier to note that the probability is simply the ratio of the length of (c,d) to the length of (a,b).



138

Uniform RV: Expectation

- Intuitively we anticipate E(X)=(a+b)/2, the midpoint of the interval. This turns out to be correct.
  - If X~uniform(a,b) we calculate

$$E(X) = \int_{-\infty}^{\infty} xf(x)dx =$$

$$\int_a^b \frac{1}{b-a} x dx = \frac{1}{2(b-a)} x^2 \Big|_a^b = \frac{b^2-a^2}{2(b-a)}$$

$$= \frac{(b+a)(b-a)}{2(b-a)} = \frac{b+a}{2}$$

139

Uniform RV: Variance

- Let X~uniform(a,b). We can find the variance of X using the shortcut formula $Var(X) = E(X^2) - \mu^2$. We proceed as follows.

$$E(X^2) = \int_a^b \frac{1}{b-a} x^2 dx = \frac{x^3}{3(b-a)}\Big|_a^b = \frac{b^3-a^3}{3(b-a)}$$

$$= \frac{b^2+ab+a^2}{3}$$

140

Uniform RV: Variance

- Finally,

$$E(X^2) - \mu^2 = \frac{b^2+ab+a^2}{3} - \frac{(a+b)^2}{4}$$

$$= \frac{4b^2+4ab+4a^2}{12} - \frac{3b^2+6ab+3a^2}{12}$$

$$= \frac{b^2-2ab+a^2}{12} = \frac{(b-a)^2}{12}$$

141

Uniform RV: Application

- Uniform random variables are the correct model for choosing a number “at random” from an interval. They are also natural choices for experiments in which some event is “equally likely” to happen at any time or place within some interval.

142

Uniform RV: Example

- Examples of probabilities involving uniform random variables are generally trivial. For instance, suppose you are trying to find a short in a 3m length of wire. Supposing you have no reason to suspect one part of the wire over another, what is the probability you find the short in the last half-meter of the wire?
- Answer: 1/6.
- You can find this by integration or by the cdf, just as you can hunt squirrels with a bazooka: Doing it once to illustrate the concept may be justifiable, but your motives become suspect if you do it more often.

143

Uniform RV: Warnings

- There is no uniform distribution on an infinite interval. In particular there is no uniform distribution on the real line. Consequently there is no way to pick a real number (“any real number”) at random, making all choices “equally likely.”
- Functions of uniform random variables may not be uniform. For instance, if $X \sim \text{uniform}(0,1)$, and Y is the square of X, then Y is a non-uniform random variable. Let us find its cdf and thereby its pdf.

144

Uniform RV: Warnings

- In practice this means that we must beware of assignments like “Choose a random square no larger than 1 inch per side.” If we choose the side length randomly, then by the above calculations we see the mean area of such a square is 1/3 square inches. If, however, we choose the area randomly (still between 0 and 1, of course), then the mean area is ½ square inches. These are two different distributions and our assignment is ambiguous.

145

Exponential Random Variables

Time between random events / time till first random event ?

If a Poisson process has constant average rate ν, the mean after a time t is $\lambda = \nu t$.

What is the probability distribution for the time to the first event?

⇒ **Exponential distribution**

Poisson - *Discrete* distribution: P(number of events)

Exponential - *Continuous* distribution: P(time till first event)

146

Exponential Random Variables: Pdf

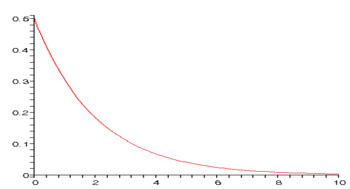
- Let λ be a positive real number. We write X~exponential(λ) and say that X is an exponential random variable with parameter λ if the pdf of X is

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

147

Exponential Random Variables: Pdf

- A look at the graph of the pdf is informative. Here is a graph for λ=0.5. Note that it has the same shape as every exponential graph with negative exponent (exponential decay). The tail shrinks to 0 quickly enough to make the area under the curve equal 1. Later we will see that the expected value of an exponential random variable is 1/λ (in this case 2).



148

Exponential Random Variables: Pdf

- A simple integration shows that the area under f is 1:

$$\int_{-\infty}^{\infty} \lambda e^{-\lambda x} dx = \lim_{t\to\infty} \int_0^t \lambda e^{-\lambda x} dx = \lim_{t\to\infty} -e^{-\lambda x}\Big|_0^t$$

$$= \lim_{t\to\infty} -e^{-\lambda t} - \left(-e^{-\lambda 0}\right) = -0 + 1 = 1$$

149

Exponential Random Variables: Expectation

- So, if X~exponential(λ), then

$$E(X)=\int_0^\infty x\lambda e^{-\lambda x}dx=\lim_{t\to\infty}\int_0^t x\lambda e^{-\lambda x}dx.\text{ Let } u=x \text{ and } dv=\lambda e^{-\lambda x}dx.$$

$$\text{Then } du=dx \text{ and } v=-e^{-\lambda x}.\text{ So } E(x)=\lim_{t\to\infty}\left(uv\Big|_0^t-\int_0^t vdu\right)$$

$$=\lim_{t\to\infty}\left(xe^{-\lambda x}\Big|_0^t-\int_0^t -e^{-\lambda x}dx\right)=\lim_{t\to\infty}\left(te^{-\lambda t}-0-\left[\frac{1}{\lambda}e^{-\lambda x}\Big|_0^t\right]\right)$$

$$=\lim_{t\to\infty}\left(te^{-\lambda t}-\left[\frac{1}{\lambda}e^{-\lambda t}-\frac{1}{\lambda}e^{-\lambda 0}\right]\right)=\lim_{t\to\infty}\left(te^{-\lambda t}-\frac{1}{\lambda}e^{-\lambda t}+\frac{1}{\lambda}\right)$$

$$=0-0+\frac{1}{\lambda}=\frac{1}{\lambda}$$

150

Exponential Random Variables: Variance

- By a similar computation .

$$Var(X)=\frac{1}{\lambda^2}$$

151

Exponential Random Variables: Applications

- Exponential distributions are sometimes used to model waiting times or lifetimes. That is, they model the time until some event happens or something quits working. Of course mathematics cannot tell us that exponentials are right to describe such situation. That conclusion depends on finding data from such real-world situations and fitting it to an exponential distribution. .

152

Exponential Random Variables: Applications

- Suppose the wait time X for service at the post office has an exponential distribution with mean 3 minutes. If you enter the post office immediately behind another customer, what is the probability you wait over 5 minutes? Since E(X)=1/λ=3 minutes, then λ=1/3, so X~exponential(1/3).
- We want .

$$P(X>5)=1-P(X\le 5)=1-F(5)$$

$$=1-\left(1-e^{-\frac{1}{3}\cdot 5}\right)=e^{-\frac{5}{3}}\approx 0.189$$

153

Exponential Random Variables: Applications

- Under the same conditions, what is the probability of waiting between 2 and 4 minutes? Here we calculate .

$$P(2 \le X \le 4) = F(4) - F(2) = \left(1 - e^{-\frac{4}{3}}\right) - \left(1 - e^{-\frac{2}{3}}\right)$$

$$= e^{-\frac{2}{3}} - e^{-\frac{4}{3}} \approx 0.250$$

154

Exponential Random Variables: Applications

- The trick in the previous example of calculating

$$P(a \le X \le b) = F(b) - F(a)$$

is quite common. It is the reason the cdf is so useful in computing probabilities of continuous random variables. .

155

Exponential Random Variables: The Memoryless Property

- The exponential random variable has an astonishing property. If X~exponential(λ) represents a waiting time, then the probability of waiting a given length of time is not affected by how long you have waited already.
- That is, $P(X > a + b \mid X > a) = P(X > b)$.
- If you have already waited a minutes, the probability you wait b more minutes is the same as your initial probability of waiting b minutes. This is known as the memoryless property.

.

156

Exponential Random Variables: The Memoryless Property

- Suppose you enter the post office and have to choose one of two lines, each of which has exactly one person ahead of you. The person in the first line got there just ahead of you. The person in the second line has already been there 10 minutes. Which line should you get in so as to be served fastest? If the waiting times are exponential, it does not matter. Similarly, if you are waiting for someone to get off a pay phone and you want to calculate the probability you have to wait more than 5 minutes, it is irrelevant to your calculations to know how long the person has already been on the phone.

157

Exponential Random Variables: The Memoryless Property

- In this respect exponential random variables behave like geometric ones. If you are finding the probability that, starting now, you flip a coin four times before getting heads, it is irrelevant to know how many times the coin flipped tails beforehand.

158

Exponential Random Variables: The Poisson Connection

- Suppose that X is a Poisson random variable measuring the number of events of some sort that happen per unit time, and suppose the mean is λ, so that X~Poisson(λ). Now let Y be the time you wait until the next Poisson event. Let us find the cdf of Y.
- Clearly F(y)=0 if y is negative.

159

Exponential Random Variables: The Poisson Connection

- Now suppose y is nonnegative. Then

$$F(y) = P(Y \le y) = 1 - P(Y > y)$$

which is the probability of having 0 Poisson events over the interval [0,y]. The number of events over [0,y] is Poisson(λy), so the probability of 0 events over [0,y] is

$$\frac{(\lambda y)^0}{0!} e^{-\lambda y} = e^{-\lambda y}$$

Thus $F(y) = 1 - e^{-\lambda y}$, which is precisely the cdf for an exponential random variable with parameter λ.

160