Some common families of continuous random variables

Math 30530, Fall 2012

November 27, 2012

The Uniform random variable

- **Name**: $\text{Uniform}(a, b)$
- **When to use**: When you want to model selecting a random number in an interval, with no part of the interval favored over any other
- **Parameters**:
  - a: the start point of the interval
  - b: the end point of the interval
- **Density function**:

$$f_X(x) = \begin{cases} 0 & \text{if } x < a \\ \frac{1}{b-a} & \text{if } a \leq x \leq b \\ 0 & \text{if } x > b \end{cases}$$

- **Statistics**:
  - $\mu = E(X) = \frac{b+a}{2}$
  - $\sigma^2 = \text{Var}(X) = \frac{(b-a)^2}{12}$

The two-dimensional Uniform random variable

- **Name**: $\text{Uniform}(R)$
- **When to use**: When you want to model selecting a random point in a finite region of the plane, with no part of the region favored over any other
- **Parameter**:
  - R: the region of interest
- **What it really is**: A pair (X, Y) of random variables, X the x-coordinate of the chosen point, Y the y-coordinate
- **Density function**:

$$f_{X,Y}(x,y) = \begin{cases} 0 & \text{if } (x,y) \notin R \\ \frac{1}{\text{Area}(R)} & \text{if } (x,y) \in R \end{cases}$$

- **Statistics**: None, since it's a pair of random variables

The Poisson process

Events occur repeatedly over a period of time

- Occurrences in disjoint time intervals are independent
- Simultaneous occurrences are very rare
- The average number of occurrences per unit time is constant throughout the time period (usually denoted λ)

The Exponential random variable

- **Name**: Exponential(λ)
- **When to use**: When you are measuring the time until the first occurrence of an event, when the occurrences satisfy the conditions of the Poisson process
- **Parameter**:
  - λ: the average number of occurrences per unit time
- **Density function**:

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

- **Statistics**:
  - $\mu = E(X) = \frac{1}{\lambda}$
  - $\sigma^2 = \mathrm{Var}(X) = \frac{1}{\lambda^2}$

Uses of the Normal random variable

- Models distribution of many physical measurements
  - height
  - weight
  - ...
- Models error made by measuring instruments
- Give a good approximation to $\text{Binomial}(n, p)$ for large n and fixed p
- Models the distribution of a quantity that is the aggregate of lots of mostly independent factors or smaller quantities
- Models the distribution of the sum of independent, identically distributed random variables
- ...
- A good distribution to use when you know (roughly) the average and variance of a quantity being measured, know that the measurements fall off at the same rate on both sides of the mean, but don't know the exact distribution

The Standard Normal random variable

- **Name**: $Z = \mathcal{N}(0, 1)$
- **When to use**: When you have problems concerning the general normal, and want to use a table to calculate associated probabilities. Transformation that takes general normal X with mean μ and variance σ^2 to standard normal is

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

- **Parameters**:
  - None
- **Density function**:

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

- **Statistics**:
  - $\mu = E(Z) = 0$
  - $\sigma^2 = \mathrm{Var}(Z) = 1$

The Normal random variable

- **Name**: $\mathcal{N}(\mu, \sigma^2)$
- **When to use**: Numerous situations; see separate page
- **Relation to standard normal** Z:

$$X = \sigma Z + \mu \qquad \text{and} \qquad Z = \frac{X - \mu}{\sigma}$$

- **Parameters**:
  - μ: the average value
  - σ^2: the variance
- **Density function**:

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

- **Statistics**:
  - $\mu = E(X) = \mu$
  - $\sigma^2 = \mathrm{Var}(X) = \sigma^2$

Sums of independent normal random variables

Let $X_1, X_2, \ldots, X_n$ be independent normal random variables, each with mean μ and variance σ^2, and set

$$S_n = X_1 + X_2 + \ldots + X_n.$$

Then

$$S_n \sim \mathcal{N}(n\mu, n\sigma^2)$$

and

$$\frac{S_n - n\mu}{\sqrt{n}\sigma} \sim \mathcal{N}(0, 1) = Z$$

- Both statements are *exact*

The central limit theorem

Let $X_1, X_2, \ldots, X_n$ be independent, identically distributed random variables, each with mean μ and variance σ^2, and set

$$S_n = X_1 + X_2 + \ldots + X_n.$$

Then for large n

$$S_n \approx \mathcal{N}(n\mu, n\sigma^2)$$

and

$$\frac{S_n - n\mu}{\sqrt{n}\sigma} \approx \mathcal{N}(0,1) = Z$$

- Both statements are *approximate*
- Works for *any* starting random variable X, discrete or continuous
- An exact form of the second statement: for each $-\infty < t < \infty$,

$$\lim_{n\to\infty} \Pr\left(\frac{S_n - n\mu}{\sqrt{n}\sigma} \leq t\right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{t} e^{\frac{-x^2}{2}}\, dx$$

DeMoivre-Laplace Theorem, Poisson approximation

DeMoivre-Laplace: If $X \sim \mathrm{Binomial}(n, p)$ then

$$X \approx \mathcal{N}(np, npq)$$

Rule of thumb: ok when $np(1-p) > 10$

The continuity correction: if $X \sim \mathrm{Binomial}(n, p)$ and $Y \sim \mathcal{N}(np, npq)$,

To Calculate	Use
$\Pr(a \leq X \leq b)$	$\Pr(a - .5 \leq Y \leq b + .5)$
$\Pr(a < X < b)$	$\Pr(a + .5 \leq Y \leq b - .5)$
$\Pr(a \leq X < b)$	$\Pr(a - .5 \leq Y \leq b - .5)$
$\Pr(a < X \leq b)$	$\Pr(a + .5 \leq Y \leq b + .5)$

Use whenever central limit theorem is used to approximate the sum of independent *discrete* random variables

Poisson approximation: If $X \sim \mathrm{Poisson}(\lambda)$ then

$$X \approx \mathcal{N}(\lambda, \lambda)$$

Rule of thumb: ok when $\lambda > 10$

The Gamma random variable

- **Name**: $\text{Gamma}(r, \lambda)$
- **When to use**: When you are measuring the time until the *r*th occurrence of an event, when the occurrences satisfy the conditions of the Poisson process
- **Parameters**:
  - λ: the average number of occurrences per unit time
  - r: the number of occurrences you are waiting to see
- **Density function**:

$$f_X(x) = \begin{cases} 0 & \text{if } x < 0 \\ \frac{\lambda^r}{(r-1)!} x^{r-1} e^{-\lambda x} & \text{if } x \geq 0 \end{cases}$$

- **Statistics**:
  - $\mu = E(X) = \frac{r}{\lambda}$
  - $\sigma^2 = \text{Var}(X) = \frac{r}{\lambda^2}$