Discrete-Event Simulation:

A First Course

Section 7.1: Continuous Random Variables

Section 7.1: Continuous Random Variables

- A random variable X is *continuous* if and only if its set of possible values $\mathcal{X}$ is a *continuum*
- A continuous random variable X is uniquely determined by
  - Its set of possible values $\mathcal{X}$
  - Its *probability density function* (pdf):
    A real-valued function $f(\cdot)$ defined for each $x \in \mathcal{X}$

$$\int_a^b f(x)dx = \Pr(a \leq X \leq b)$$

By definition,

$$\int_{\mathcal{X}} f(x)dx = 1$$

Example 7.1.1

- X is *Uniform*(a, b)
  $\mathcal{X} = (a, b)$ and all values in this interval are equally likely

$$f(x) = \frac{1}{b-a} \qquad a < x < b$$

- In the continuous case,
  - $\Pr(X = x) = 0$ for *any* $x \in \mathcal{X}$
  - If $[a, b] \subseteq \mathcal{X}$,

$$\begin{aligned} \int_a^b f(x)dx &= \Pr(a \le X \le b) = \Pr(a < X \le b) \\ &= \Pr(a \le X < b) = \Pr(a < X < b) \end{aligned}$$

Cumulative Distribution Function

- The *cumulative distribution function(cdf)* of the continuous random variable X is the real-valued function $F(\cdot)$ for each $x \in \mathcal{X}$ as

$$F(x) = \Pr(X \leq x) = \int_{t \leq x} f(t)dt$$

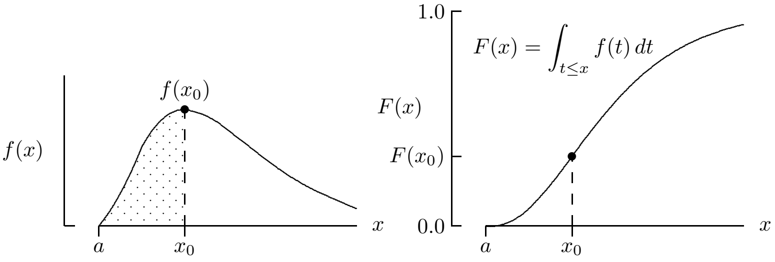
- **Example 7.1.2:** If X is *Uniform*(a, b), the cdf is

$$F(x) = \int_{t=a}^{x} \frac{1}{(b-a)} dt = \frac{x-a}{b-a} \qquad a < x < b$$

- In special case where U is *Uniform*$(0, 1)$, the cdf is

$$F(u) = \Pr(U \leq u) = u \qquad 0 \leq u \leq 1$$

Relationship between pdfs and cdfs



- Shaded area in pdf graph equals $F(x_0)$

More on cdfs

- The cdf is strictly monotone increasing:
  if $x_1 < x_2$, then $F(x_1) < F(x_2)$
- The cdf is bounded between 0.0 and 1.0
- The cdf can be obtained from the pdf by integration
  The pdf can be obtained from the cdf by differentiation as

$$f(x) = \frac{d}{dx}F(x) \qquad x \in \mathcal{X}$$

- A continuous random variable model can be specified by $\mathcal{X}$ and either the pdf or the cdf

Example 7.1.3: *Exponential*(μ)

- $X = -\mu \ln(1 - U)$ where U is *Uniform*$(0, 1)$
- The cdf of X is

$$\begin{aligned} F(x) = \Pr(X \le x) &= \Pr(-\mu \ln(1 - U) \le x) \\ &= \Pr(1 - U \ge \exp(-x/\mu)) \\ &= \Pr(U \le 1 - \exp(-x/\mu)) \\ &= 1 - \exp(-x/\mu) \end{aligned}$$

- The pdf of X is

$$f(x) = \frac{d}{dx}F(x) = \frac{d}{dx}(1 - \exp(-x/\mu)) = \frac{1}{\mu}\exp(-x/\mu) \qquad x > 0$$

Mean and Standard Deviation

- The *mean* μ of the continuous random variable X is

$$\mu = \int_x x f(x) dx$$

- The corresponding *standard deviation* σ is

$$\sigma = \sqrt{\int_x (x-\mu)^2 f(x) dx} \qquad \text{or} \qquad \sigma = \sqrt{\left(\int_x x^2 f(x) dx\right) - \mu^2}$$

- The *variance* is σ^2

Examples

- If X is *Uniform*(a, b)

$$\mu = \frac{a+b}{2} \qquad \text{and} \qquad \sigma = \frac{b-a}{\sqrt{12}}$$

- If X is *Exponential*(μ),

$$\int_x x f(x) dx = \int_0^\infty \frac{x}{\mu} \exp(-x/\mu) dx = \mu \int_0^\infty t \exp(-t) dt = \cdots = \mu$$

$$\sigma^2 = \left(\int_0^\infty \frac{x^2}{\mu} \exp(-x/\mu) dx \right) - \mu^2 = \cdots = \mu^2$$

Expected Value

- The mean of a continuous random variable is also known as the *expected value*
- The expected value of the continuous random variable X is

$$\mu = E[X] = \int_x x f(x) dx$$

- The variance is the expected value of $(X - \mu)^2$

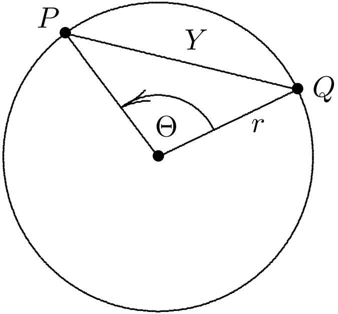
$$\sigma^2 = E[(X - \mu)^2] = \int_x (x - \mu)^2 f(x) dx$$

- In general, if $Y = g(X)$, the expected value of Y is

$$E[Y] = E[g(X)] = \int_x g(x) f(x) dx$$

Example 7.1.6

- A circle of radius r and a fixed point Q on the circumference
- P is selected *at random* on the circumference
- Let the random variable Y be the distance of the line segment joining P and Q



$$Y = 2r\sin(\Theta/2)$$

Example 7.1.6 ctd.

- If Θ is *Uniform*$(0, 2\pi)$, the pdf of Θ is $f(\theta) = 1/2\pi$
- The expected length of Y is

$$E[Y] = \int_0^{2\pi} 2r\sin(\theta/2)f(\theta)d\theta = \int_0^{2\pi} \frac{2r\sin(\theta/2)}{2\pi}d\theta = \cdots = \frac{4r}{\pi}$$

- Y is not *Uniform*$(0, 2r)$; otherwise, $E[Y]$ would be r.

Example 7.1.7

- If continuous random variable $Y = aX + b$ for constants a and b,

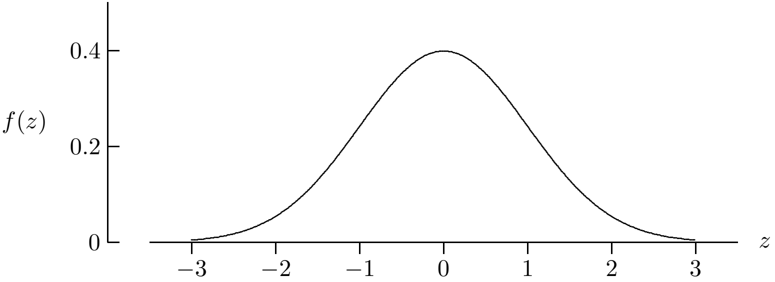
$$E[Y] = E[aX + b] = aE[X] + b$$

Continuous Random Variable Models

- **Standard Normal Random Variable**

  Z is *Normal*$(0,1)$ if and only if the set of all possible values is $\mathcal{Z} = (-\infty, \infty)$ and the pdf is

$$f(z) = \frac{1}{\sqrt{2\pi}} \exp(-z^2/2) \qquad -\infty < z < \infty$$



Standard Normal Random Variable

- If Z is *Normal*$(0, 1)$, Z is "standardized"

  The mean is

$$\mu = \int_{-\infty}^{\infty} zf(z)dz = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} z \exp(-z^2/2)dz = \cdots = 0$$

  The variance is

$$\sigma^2 = \int_{-\infty}^{\infty} (z-\mu)^2 f(z)dz = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} z^2 \exp(-z^2/2)dz = \cdots = 1$$

- The cdf is

$$F(z) = \int_{-\infty}^{z} f(t)dt = \Phi(z) \qquad -\infty < z < \infty$$

Standard Normal cdf

- $\Phi(\cdot)$ is defined as

$$\Phi(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} \exp(-t^2/2)dt \qquad -\infty < z < \infty$$

- No closed-form expression for $\Phi(z)$

$$\Phi(z) = \begin{cases} \dfrac{1 + P(1/2, z^2/2)}{2} & z \geq 0 \\ 1 - \Phi(z) & z < 0 \end{cases}$$

  $P(a, x)$ is an incomplete gamma function (see Appendix D)
- Function $\Phi(z)$ is available in `rvms` as `cdfNormal(0.0, 1.0, z)`

Scaling and Shifting

- Suppose X is a random variable with mean μ and standard deviation σ
- Define random variable $X' = aX + b$ for constants a, b
- The mean μ' and standard deviation σ' of X' are

$$\mu' = E[X'] = E[aX + b] = aE[X] + b = a\mu + b$$

$$(\sigma')^2 = E[(X' - \mu')^2] = E[(aX - a\mu)^2] = a^2 E[(X - \mu)^2] = a^2\sigma^2$$

Therefore,

$$\mu' = a\mu + b \qquad \text{and} \qquad \sigma' = |a|\sigma$$

Example 7.1.8

- Suppose Z is a random variable with mean 0 and standard deviation 1
- Construct a new random variable X with *specified* mean μ and standard deviation σ
- Define $X = \sigma Z + \mu$
- $E[X] = \sigma E[Z] + \mu = \mu$
- $E[(X - \mu)^2] = E[\sigma^2 Z^2] = \sigma^2 E[Z^2] = \sigma^2$

Normal Random Variable

- The continuous random variable X is *Normal*(μ, σ) if and only if

$$X = \sigma Z + \mu$$

  where $\sigma > 0$ and Z is *Normal*$(0, 1)$
- The mean of X is μ and the standard deviation is σ
- *Normal*(μ, σ) is constructed from *Normal*$(0, 1)$
  - by "shifting" the mean from 0 to μ via the addition of μ
  - by "scaling" the standard deviation from 1 to σ via multiplication by σ

cdf of Normal Random Variable

- The cdf of a *Normal*(μ, σ)

$$F(x) = \Pr(X \le x) = \Pr(\sigma Z + \mu \le x) = \Pr(Z \le (x - \mu)/\sigma)$$

so that

$$F(x) = \Phi\left(\frac{x-\mu}{\sigma}\right) \qquad -\infty < x < \infty$$

where $\Phi(\cdot)$ is the cdf of *Normal*$(0, 1)$

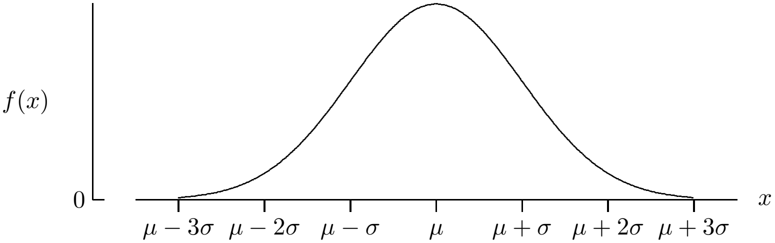
pdf of Normal Random Variable

- Because

$$\frac{d}{dz}\Phi(z) = \frac{1}{\sqrt{2\pi}}\exp(-z^2/2) \qquad -\infty < z < \infty$$

the pdf of *Normal*(μ, σ) is

$$f(x) = \frac{dF(x)}{dx} = \frac{d}{dx}\Phi\left(\frac{x-\mu}{\sigma}\right) = \cdots = \frac{1}{\sigma\sqrt{2\pi}}\exp(-(x-\mu)^2/2\sigma^2)$$



Some Properties of Normal Random Variables

- Sums of *iid* random variables approach the normal distribution
- *Normal*(μ, σ) is sometimes called a *Gaussian* random variable
- The 68-95-99.73 rule
  - Area under pdf between $\mu - \sigma$ and $\mu + \sigma$ is about 0.68
  - Area under pdf between $\mu - 2\sigma$ and $\mu + 2\sigma$ is about 0.95
  - Area under pdf between $\mu - 3\sigma$ and $\mu + 3\sigma$ is about 0.9973
- The pdf has inflection points at $\mu \pm \sigma$
- Common notation for *Normal*(μ, σ) is $N(\mu, \sigma^2)$
- Support is $\mathcal{X} = \{x \mid -\infty \leq x \leq \infty\}$
  - Usually not appropriate for simulation unless modified to produce only positive values

Lognormal Random Variable

- The continuous random variable X is *Lognormal*(a, b) if and only if

$$X = \exp(a + bZ)$$

  where Z is *Normal*$(0, 1)$ and $b > 0$
- *Lognormal*(a, b) is also based on transforming *Normal*$(0, 1)$
  - The transformation is non-linear

cdf of Lognormal Random Variable

- The cdf of a *Lognormal*(a, b)

  $$F(x) = \Pr(X \leq x) = \Pr(\exp(a+bZ) \leq x) = \Pr(a+bZ \leq \ln(x))$$

  so that

  $$F(x) = \Pr(Z \leq (\ln(x) - a)/b) = \Phi\left(\frac{\ln(x) - a}{b}\right) \qquad x > 0$$

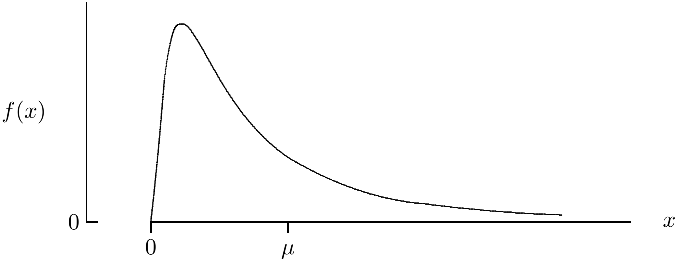
  where $\Phi(\cdot)$ is the cdf of *Normal*$(0, 1)$

pdf of Lognormal Random Variable

- The pdf of *Lognormal*(a, b) is

$$f(x) = \frac{dF(x)}{dx} = \cdots = \frac{1}{bx\sqrt{2\pi}} \exp(-(\ln(x)-a)^2/2b^2) \qquad x > 0$$

$$(a, b) = (-0.5, 1.0)$$



- $\mu = \exp(a + b^2/2)$ Above, $\mu = 1.0$
- $\sigma = \exp(a + b^2/2)\sqrt{\exp(b^2) - 1}$ Above, $\sigma \simeq 1.31$

Erlang Random Variable

- *Uniform*(a, b) is the continuous analog of *Equilikely*(a, b)
- *Exponential*(μ) is the continuous analog of *Geometric*(p)
- *Pascal*(n, p) is the sum of n *iid Geometric*(p)
- What is the continuous analog of *Pascal*(n, p)?

  The continuous random variable X is *Erlang*(n, b) if and only if

  $$X = X_1 + X_2 + \cdots + X_n$$

  where $X_1, X_2, \cdots, X_n$ are *iid Exponential*(b) random variables

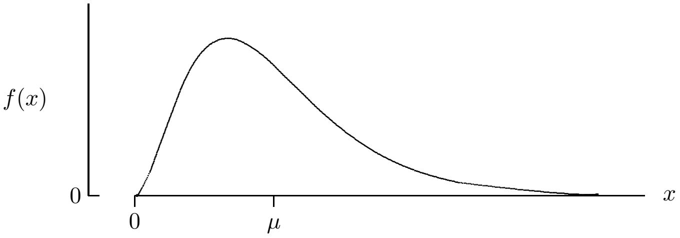
0-13-142917-5

pdf of Erlang Random Variable

- The pdf of *Erlang*(n, b) is

$$f(x) = \frac{1}{b(n-1)!}(x/b)^{n-1}\exp(-x/b) \qquad x > 0$$

$$(n, b) = (3, 1.0)$$



- For $(n, b) = (3, 1.0)$, $\mu = 3.0$ and $\sigma \simeq 1.732$

cdf of Erlang Random Variable

- The corresponding cdf is

$$F(x) = \int_0^x f(t)dt = P(n, x/b) \qquad x > 0$$

  Incomplete gamma function (see Appendix D)

- $\mu = nb$
- $\sigma = \sqrt{n}b$

Chisquare And Student Random Variables

- $Chisquare(n)$ and $Student(n)$ are commonly used for statistical inference
- Defined in section 7.2