EC381/MN308
Probability and Some Statistics

Yannis Paschalidis

yannisp@bu.edu, http://ionia.bu.edu/

Dept. of Manufacturing Engineering

Dept. of Electrical and Computer Engineering

Center for Information and Systems Engineering

Lecture 8 - Outline

1. Families of Continuous Random Variables

3.4 Families of Continuous RVs

1) Uniform
2) Exponential
3) Erlang (gamma)
4) Gaussian (Normal)

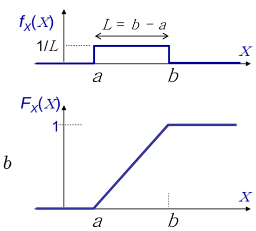
These four families of continuous RVs describe the outcome of common experiments.

Members of a given family differ only by the value(s) of the (one or two) parameters of the family.

Many other families of continuous RVs exist as well.

1) Uniform Random Variables

X is uniformly distributed over $[a, b]$



PDF: $f_X(x) = 1/L$ for $a \le x \le b$

CDF: $F_X(x) = \begin{cases} 0 & x < a \\ \frac{x-a}{b-a} & a \le x \le b \\ 1 & x > b \end{cases}$

Mean: $E[X] = (a + b)/2$ = *midpoint* of interval

Variance: $\mathrm{Var}[X] = (b - a)^2/12 = L^2/12$

Applications: To model complete lack of knowledge, or complete indifference within a range of values.

Compare with discrete uniform PMF studied in Chapter 2.

Example

Signal with unknown phase angle Θ, e.g., a sinusoidal signal arriving with an unknown time delay

PDF $f_\Theta(u) = \dfrac{1}{2\pi}$

$E[\Theta] = \pi$

$$\mathrm{Var}[\Theta] = \int_0^{2\pi} u^2 \frac{1}{2\pi} du - \pi^2$$
$$= \frac{1}{6\pi} 8\pi^3 - \pi^2 = \frac{1}{3}\pi^2 = \frac{(2\pi)^2}{12}$$

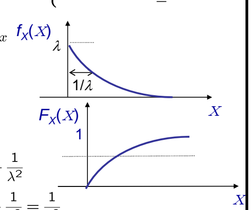
2) Exponential Random Variables

X is an exponential random variable with a single parameter $\lambda > 0$, called the scale parameter

PDF: $f_X(x) = \begin{cases} \lambda e^{-\lambda x} & x \ge 0 \\ 0 & x < 0 \end{cases}$ CDF: $F_X(x) = \begin{cases} 0 & x < 0 \\ 1 - e^{-\lambda x} & x \ge 0 \end{cases}$

Mean: $E[X] = \int_0^\infty x\lambda e^{-\lambda x} dx = \int_0^\infty -x de^{-\lambda x}$

Integrate by parts: $v\,du = uv - u\,dv$ $= -xe^{-\lambda x}\Big|_0^\infty + \int_0^\infty e^{-\lambda x} dx = \frac{1}{\lambda}$



Variance: $\mathrm{Var}[X] = \int_0^\infty x^2 \lambda e^{-\lambda x} dx - \frac{1}{\lambda^2}$

Integrate by parts
$$= -x^2 e^{-\lambda x}\Big|_0^\infty + \int_0^\infty 2xe^{-\lambda x} dx - \frac{1}{\lambda^2}$$
$$= -2\frac{x}{\lambda} e^{-\lambda x}\Big|_0^\infty + \frac{2}{\lambda}\int_0^\infty e^{-\lambda x} dx - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}$$

Applications: Studies of waiting times and service times.
Continuous version of "delay" model: analogous to the discrete geometric RV with parameter p and mean $1/p$, which is the waiting time in discrete units.

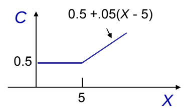
Example

(1) The duration of a phone call X is an exponentially distributed RV, with parameter λ = 0.1 per min. Determine the expected duration of a call.

(2) The charge for the call is \$0.50 for the first five minutes plus \$0.05 per minute thereafter. Determine the mean and variance of the cost per call.

Expected duration of call: $E[X] = \frac{1}{\lambda} = 10$ min

Cost C of call:



Mean cost:

$$E[C] = E[0.5 + 0.05(X-5)I\{X>5\}]$$
$$= 0.5 + 0.05\int_5^\infty (x-5)0.1e^{-0.1x}dx$$
$$= 0.5 + 0.05\int_5^\infty (x-5)0.1e^{-0.1(x-5)}e^{-.5}dx \text{ Change}: y = x-5$$
$$= 0.5 + 0.05e^{-0.5}\cdot 10 = 0.5(1+e^{-0.5}) \approx 0.803$$

Variance of X: $\text{Var}[X] = \frac{1}{\lambda^2} = 100 \quad E[X^2] = \frac{2}{\lambda^2} = 200$

$\sigma_X = 10$ min

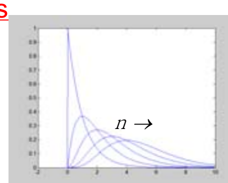
Variance of C

$$E[C^2] = E[(0.5 + 0.05(X-5)I\{X>5\})^2]$$
$$= 0.25 + E[0.05(X-5)I\{X>5\}] + E[0.05^2(X-5)^2 I\{X>5\}]$$
$$= 0.25 + .303 + 0.05^2\int_5^\infty (x-5)^2 0.1e^{-0.1(x-5)}e^{-.5}dx$$
$$= 0.25 + 0.303 + 0.05^2 e^{-0.5}\cdot 200$$
$$= 0.25 + 0.303 + .5e^{-0.5} \approx 0.856$$
$$Var[C] = E[C^2] - E[C]^2 = 0.856 - 0.803^2 = 0.21$$
$$\sigma_C = \$0.46$$

3) Erlang Random Variables

(Also called gamma random variables)

PDF: $f_X(x) = \begin{cases} \frac{\lambda^n x^{n-1}}{(n-1)!}e^{-\lambda x} & x \geq 0 \\ 0 & x < 0 \end{cases}$



Reduces to exponential RV for n = 1

Mean: $E[X] = n/\lambda$ (n times that of the exponential RV)

Variance: $\text{Var}[X] = n/\lambda^2$ (n times that of the exponential RV)

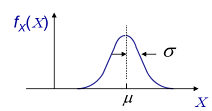
Applications: Erlang with parameters (n, λ) corresponds to the sum of n independent exponential random variables with parameter λ.

Delay after n stages, each of which has an exponential-random-variable delay with parameter λ: Application to **communication networks!**

Agner Erlang (1878-1929), a Danish mathematician, studied telephone traffic and is the father of queuing theory.

4) Gaussian (Normal) Random Variables

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



The familiar bell-shaped curve

Mean: $E[X] = \mu$

Variance: $\text{Var}[X] = \sigma^2$

We often write $X \sim N(\mu, \sigma^2)$ or $X \sim N(\mu, \sigma^2)$ or use the phrase "X is $N(\mu, \sigma^2)$" to say that X is Gaussian with mean μ and variance σ^2

Karl Friedrich Gauss (1777-1855), one of the greatest mathematicians of all time, used this PDF, named after him, to study errors in measurements.

Applications:

- Many sets of data gathered from a variety of physical phenomena fit the Gaussian (or normal) distribution
- The central limit theorem asserts that if many "small and random" causes produce a net effect, then that effect can be modeled as a normal random variable (we will study this later).

  Example: Noise voltages in electrical circuits, such as the voltage measured across a resistor at a temperature T, have normal distributions. The noise voltage is the net result of the electric fields created by many electrons at random positions: each charge has a small effect on the voltage.

Henri Poincaré (1908) observed with respect to the central limit theorem: All the world believes it firmly because the mathematicians imagine that it is a fact of observation and the observers imagine that it is a theorem of mathematics.

The central limit theorem is not *always* valid, however; families of stable distributions.

- Linear Transformations of Gaussian RVs are also Gaussian

If X is a Gaussian RV with mean μ_X and variance σ_X^2, and a, b are constants,

then $Y = aX + b$ is also Gaussian with mean & variance:

$\mu_Y = a\mu_X + b$

$\sigma_Y^2 = a^2\sigma_X^2$

Proof: later

These relations apply for any RV

Also, linear functions of multiple Gaussian RVs are Gaussian RVs.

Also, the output of a linear system is a Gaussian random variable if its input is a Gaussian random variable:

GIGO: Gaussian In $\Rightarrow$ Gaussian Out

This is another reason for the ubiquity of the Gaussian RV

Isn't everything "normal"?

- Many people seem to think that all data sets must be "normally" distributed

  …and go through lots of contortions to make the data conform!
- Example: Raw SAT and GRE scores are "curved" to get a N(500, 100^2) distribution

  A nonlinear transformation is applied to raw scores to ensure that the histogram of the result looks like a N(500, 100^2) PDF
  - Histogram: a plot of relative frequency of given scores

- Students in a class are often graded on a Bell curve

  Scores are warped to fit Bell curve. e.g., top 10% A's, 35% B's, 35% C's, 20% D's and F's

  We don't believe in warping to fit Bell curves, however …

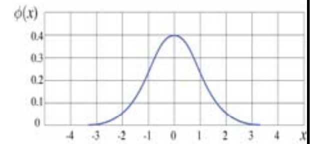
Unit or standard Gaussian RV

= Gaussian RV with zero mean and unit variance $X \sim N(0, 1)$

PDF:

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

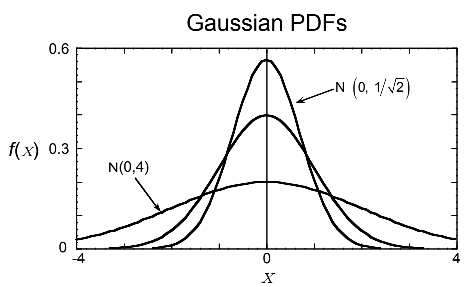
Peak at 0 has value $\frac{1}{\sqrt{2\pi}}$



Can write PDF of an arbitrary Gaussian RV with mean μ and variance σ^2 in terms of $\phi(x)$ by shifting and scaling:

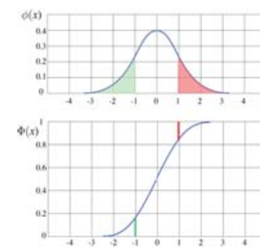
$$f_X(x) = \frac{1}{\sigma}\phi(\frac{x-\mu}{\sigma})$$

- shift to the right by μ,
- stretch laterally by a factor σ
- compress vertically by a factor $1/\sigma$ to preserve area



The CDF of the Standard Gaussian RV $X \sim N(0, 1)$

$$F_X(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} du \equiv \Phi(x)$$



$\Phi(x)$ cannot be expressed in terms of simple functions of x (e.g. *exp*, *ln*, *sin*), but the function is tabulated

$\Phi(-x)$ = green area

$1 - \Phi(x)$ = red area

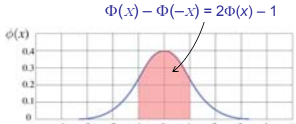
$Q(x) = 1 - \Phi(x)$ = complementary unit Gaussian CDF function (used in many applications such as communication theory)

$\Phi(x)$ is tabulated only for $x \geq 0$

$\Phi(-x) = 1 - \Phi(x)$

$\Phi(x) - \Phi(-x) = 2\Phi(x) - 1$

$\Phi(x) - \Phi(-x) = 2\Phi(x) - 1$



Red area calculated from tabulated values using $\Phi(x) - \Phi(-x) = 2\Phi(x) - 1$:

Tabulated:	
$\Phi(1) = 0.8413$	$\Phi(1) - \Phi(-1) = 0.6826$
$\Phi(2) = 0.9772$	$\Phi(2) - \Phi(-2) = 0.9544$
$\Phi(3) = 0.9987$	$\Phi(3) - \Phi(-3) = 0.9974$

$\Phi(1.96) - \Phi(-1.96) = 0.95$

Almost all (99.74%) of the probability mass lies in the interval [–3, 3] .

Computing with $\Phi(x)$

For an arbitrary Gaussian RV with mean μ and variance σ^2, i.e., $X \sim N(\mu, \sigma^2)$

$F_X(x) = \Phi(\frac{X-\mu}{\sigma})$ i.e., the CDF is a shifted and scaled version of $\Phi(x)$

Example

For a Gaussian RV X with $\mu = 1$ and $\sigma = 2$, determine the probability that X lies between -1 and 3.

$P[-1 < X < 3] = F_X(3) - F_X(-1) = F_X(3) - [1 - F_X(1)]$

$= F_X(3) + F_X(1) - 1$

$= \Phi((3-\mu)/\sigma) + \Phi((1-\mu)/\sigma) - 1$

$= \Phi((3-1)/2) + \Phi((1-1)/2) - 1 = \Phi(1) + \Phi(0) - 1$

$= 0.8413 + 0.5 - 1 = 0.3413$