Math Modeling Day 13.

* So far: two different Stochastic Models:

  * Squirrels 50/50 chance of survival
  * Monte Carlo — Gambling game.

For these cases we used:

<u>BINOMIAL DISTRIBUTION</u> — when there are only two possible outcomes
SUCCESS OR FALIURE.

A Bernoulli Trial — when you simulate a series of random events that can take only one of two values.

The outcome: $f(x) = p^x q^{1-x}$ $\quad x = 0$ or 1

p = prob of success $\quad q$ = prob of faliure.

The Binomial Distribution is what happens when multiple Bernoulli Trials are performed

the probability of success in k trials $\quad f(k) = \binom{n}{k} p^k q^{n-k}$ $\quad k = 0, 1, 2, \ldots$

$\binom{n}{k}$ = "n choose k" $= \dfrac{n!}{(n-k)!\,k!}$ combination # or in a set of n-elements the # of distinct k elem[ent] sub sets.

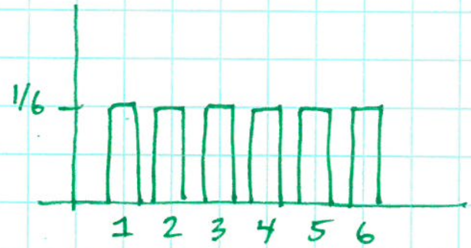
FOR US — we just simulated the Bernoulli Trials using rand().

But what about more complicated situations? For example, rainfall... This is not a flip of the coin! Rolling Dice? More than two sides.

Today... other useful distributions.

UNIFORM DISTRIBUTION — when outcomes are equally likely
DISCRETE and CONTINUOUS

in the DISCRETE CASE



ROLL OF DICE

NOTE: for flat data

Programming:
Use rand() to generate a random #
then
if rand() < $1/6$ we roll 1
if $1/6 <$ rand() $< 1/3$ we roll 2
⋮

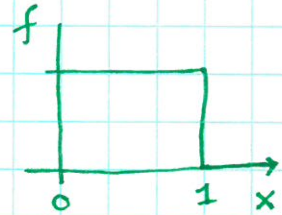
But wait .... what does rand() do — it generates a random # between 0 and 1 each w/ same probability.

CONTINUOUS CASE

$$f(x) = \begin{cases} 1/(\beta-\alpha) & \alpha < x < \beta \\ 0 & \end{cases}$$

rand() samples from

$$f(x) = \begin{cases} 1 & 0 < x < 1 \\ 0 & \end{cases}$$

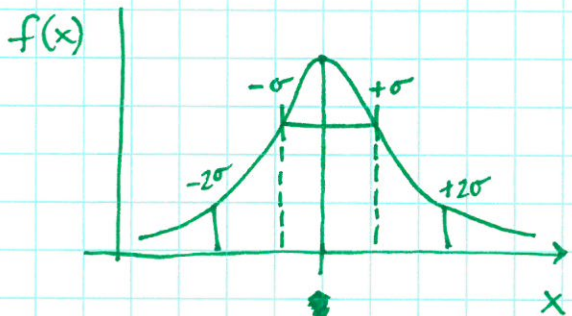


"Sampling" from a distribution means choosing values with the correct probability. In other words here we will never get a 2 but we will get between 0 and 1 with equal probability.

THE NORMAL DISTRIBUTION — Use when some values are more likely than others.

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}}\, e^{\left(\frac{-(x-\mu)^2}{2\sigma^2}\right)}$$

where: μ = mean or average
σ = stdev.
σ^2 = varriance.



"Bell curve"

Data that is normally distributed "matches" this curve.

μ – mean is most likely

TO BE SURE: 68 – 95 – 99 Rule:

68% of all your data observations fall with in ONE Standard Deviation from μ
95% of all data observations fall with in TWO Standard Deviations
99% should fall with in THREE Standard Deviations.

THE GAMMA DISTRIBUTION — Use when some outcomes are more likely than others but there is skew in the data.

$$f(x) = \frac{x^{\alpha-1}\, e^{-x\beta}}{(1/\beta)^{\alpha}\,\Gamma(\alpha)}$$

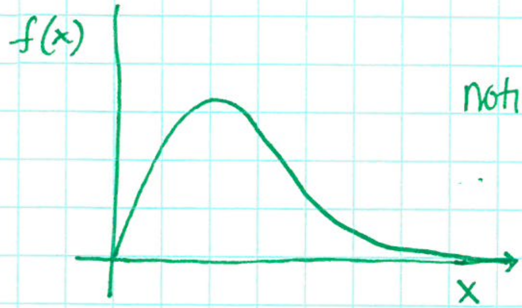
here $\Gamma(\alpha)$ is the gamma function

here $\alpha = \frac{\mu^2}{\sigma^2}$ $\beta = \frac{\sigma^2}{\mu}$

$$\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} e^{-x}\, dx$$

on $\Gamma(\alpha) = (\alpha-1)!$ for discrete values.

if $\alpha = 2$ and $\beta = \frac{1}{2}$



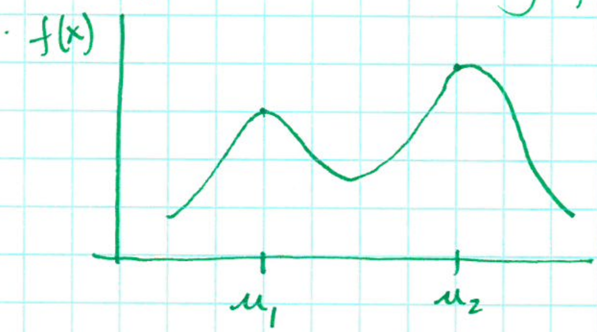
notice how it is shifted.

NOTE – when $\alpha = 1$ this becomes the exponential distribution.

<u>BIMODAL DISTRIBUTIONS</u> – when the data has two or more values that seem more likely.

$$f(x) = p\, g_1(x) + (1-p)\, g_2(x)$$

$g_1(x)$ and $g_2(x)$ are two distributions.
p is the mixing parameter.



HOW DO WE SAMPLE THES DISTRIBUTIONS or GENERATE DATA.

in a spreadsheet:

= NORMINV(rand() , mean , standard dev)

= GAMMAINV(rand() , α , β)

BIMODAL: = if(rand() < p , sample dist 1 , sample dist 2)

NOTE: There can be issues with these functions – if you go on to professional stochastic modeling YOU MUST TAKE MORE STATS / PROBABILITY !!!

Example Rainfall Data. xls.

#1 Analyze Real Data.

STATS

HISTOGRAM

#2 Choose a Distribution

UNIFORM, GAMMA

NORMAL, BIMODAL

#3 Build the Model.