STAT 280 November 19, 2001

Ch 8: Sums of Random Variables

(8.1-8.3 should be mostly review: covered already in Stat 270 and earlier in this course).

Thm 8.1-1 Sum of independent normals is normal

Thm 8.2-1 Sum of independent and identically distributed (iid) random variables is approx normal

Notes for both theorems:

1. Proof uses MGFs
2. Works for averages too
3. Results about mean and SD of sum (or average) may be verified from first principles

Also, you can verify empirically (by simulation) that Thm 8.2-1 also works for a large class of cases when the variables are not independent, and when the variables are unequally weighted.

8.3 Confidence Intervals - these are a direct consequence of the above theorems (and not important for this course).

8.4 Random sums of IID RVs.

Let N(t) be Poisson Process

Let Y_i be iid RVs i = 1,2,3,…

Let X(t) = $\sum_{i=1}^{N(t)} Y_i$. For each t, X(t) is a random sum of RVs. The ordered set {X(t): t>0} is called a Compound Poisson Process.

Mean and SD of X(t) (as function of t) can be expressed in terms of Poisson mean m and mean μ and SD σ of Y_i:

$E(X(t)) = mt\mu$ and $SD(X(t)) = [mt(\mu^2 + \sigma^2)]^{1/2}$

Insurance example: claims average $10,000 with SD of $15,000. Average number of claims per year is 25.

Mean and SD of annual payout is

Mean = $250,000 and, working in thousands of dollars,

SD = $[25(100 + 225)]^{1/2}$ = 90.139 or $90,139

So, IF we assume normality of X(t), company would need to be prepared for payouts of from as little as $70,000 to as much as $430,000.

But IS the distribution normal or approx normal?

Can't tell unless distribution of payout size is known.

CLT does not quite apply. Try simulation with an assumed distribution shape, and having given mean and SD.

Suppose payouts are Gamma with mean $10,000 and SD $15,000. Work in thousands, mean=10 SD=15.

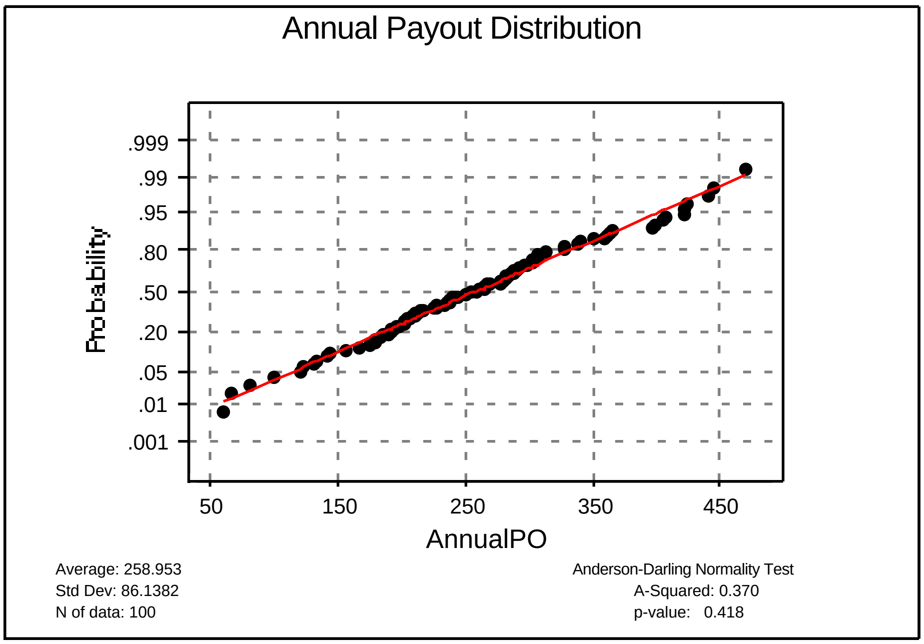
Look at mean and SD formula for gamma. Use formulas on p 252 to get $\alpha = 100/225 = .444$ $\beta = 225/10 = 22.5$.

Use these in a simulation:

```
Gmacro
cPoisson              # name of macro
brief 0               # suppress output during do-loop
do k3=1:100            # generate 100 1-year experiences
rand 1 c1;             # how many claims in a year?
poisson 25.
let k1=c1(1)           # capture number of claims in a constant (syntax
need)
rand k1 c2;           #simulate that many claims
gamma .444 22.5.   # claim distributon is gamma with mean 10 and SD
15
sum c2 k2              # add up annual claims
let c3(k3)=k2          # save result for next year's simulation
enddo
brief 2               # allow output
name c3 'AnnualPO'  # show dotplot of result
dotpl c3               # can also do normal plot to check normality.
Endmacro
```

In the following dotplot, Annual Payout is in $,000

```
                                 .
                    :    .  . :: :
                 : :::. .:. : :::.: .   .       .
   : .  . :..:. .::::::::::::.::::: :...:.   :: : ..  .
-----+---------+---------+---------+---------+---------+-AnnualPO
    80       160       240       320       400       480
```



Looks like CLT DOES work with random sums!

A little primer on the Beta distribution:

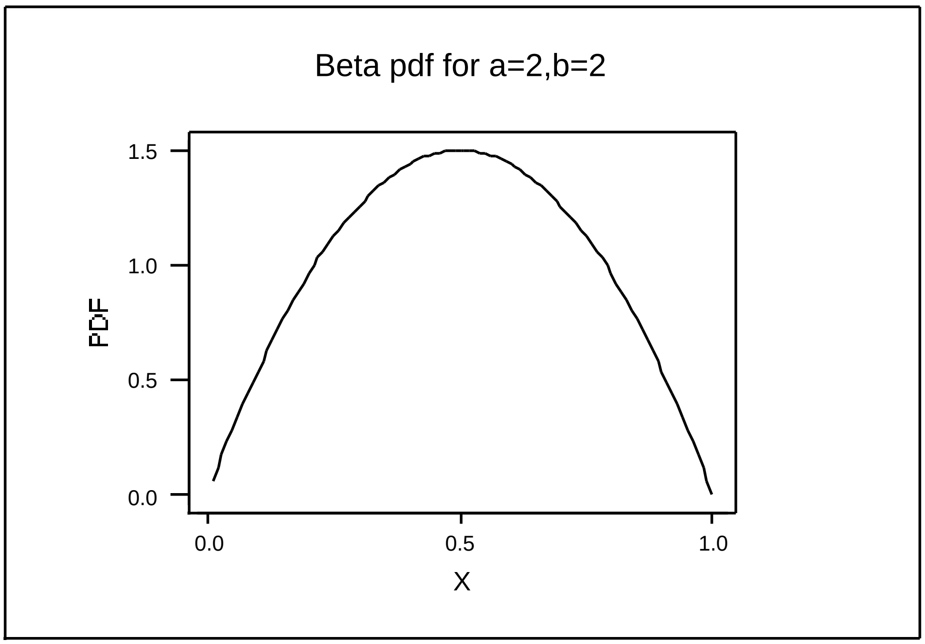
Pdf is c $x^{a-1}(1-x)^{b-1}$ where $c=\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}$

Mean is a/(a+b). SD is square root of $ab/[(a+b+1)(a+b)^2]$

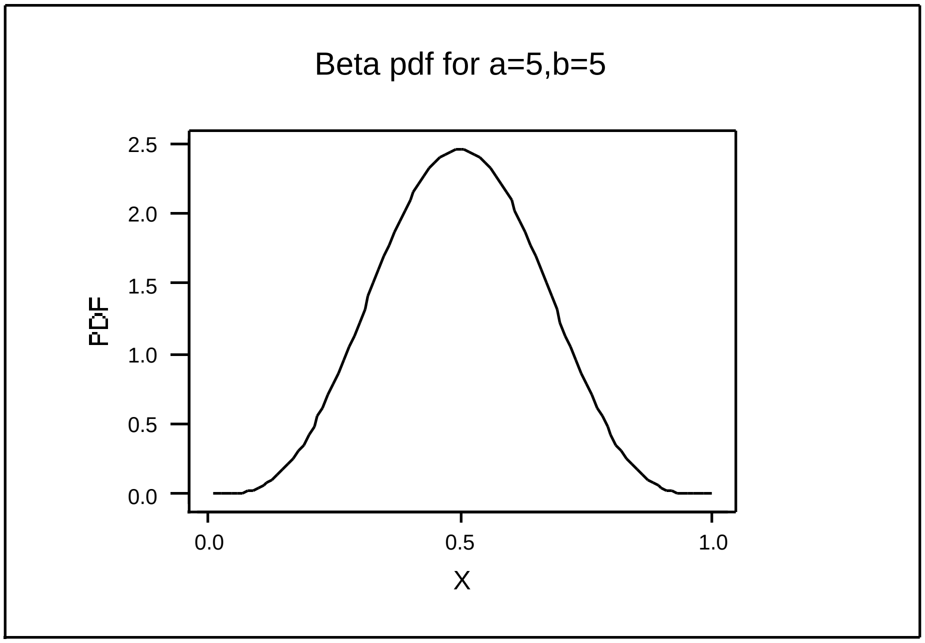
If a=1, b=1, this Beta is U(0,1)

If a=b >1, the Beta is unimodal and symmetrical
If a and b are both very large, the Beta is almost deterministic at a particular value in (0,1)

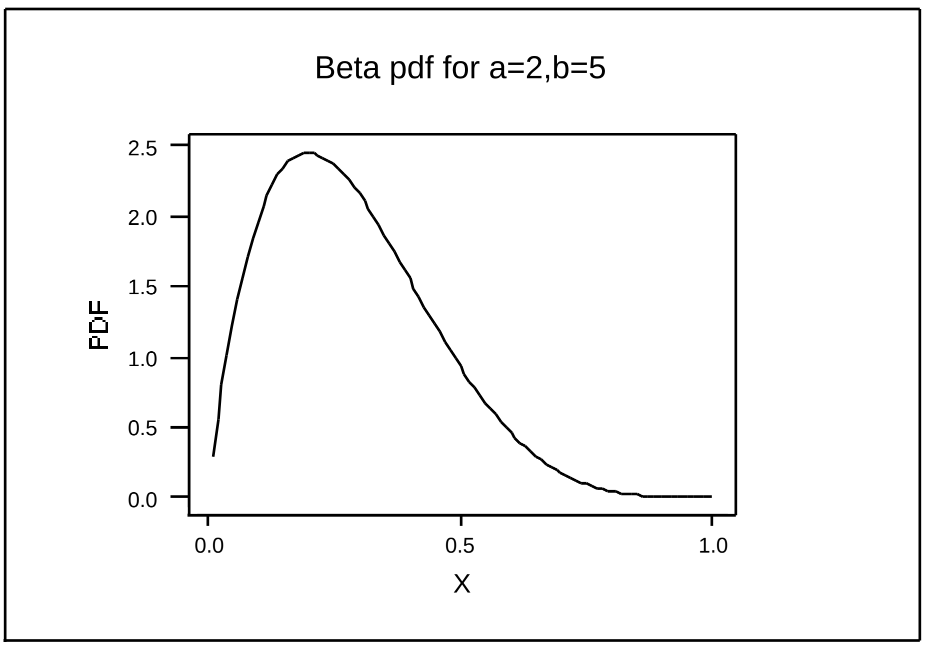
Here is a=2,b=2



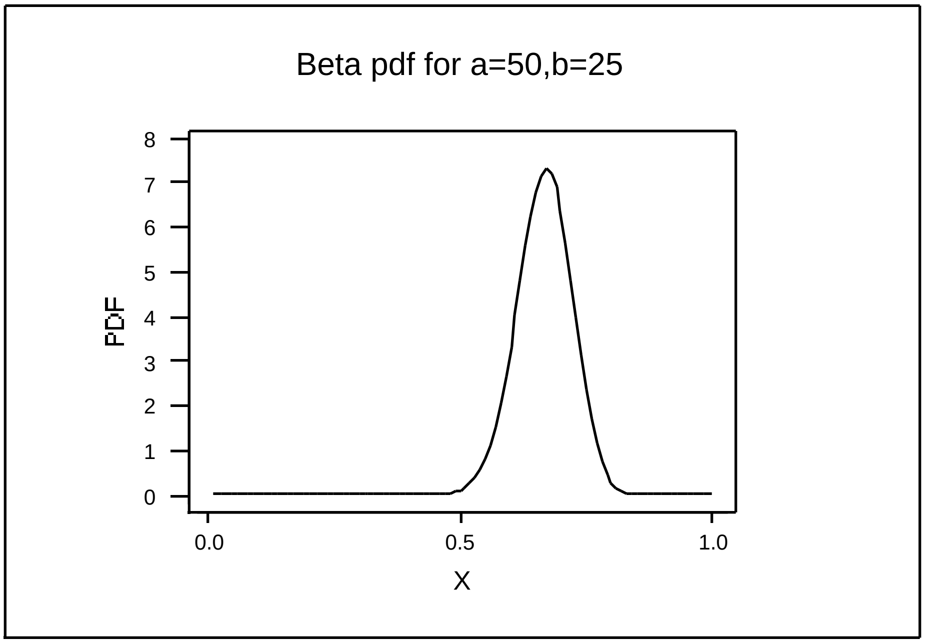
And here it is for a=5, b=5



And for a=2, b=5



And for a=50, b=25



The Beta models any RV with a fixed interval of support since we can transform X=Beta to (L,U) with L+(U-L)*X

It is useful for modeling a random probability (on (0,1))

How to get these graphs of the Beta PDF?

```
set c1
DATA> 0:100
DATA> end
MTB > let c1=c1/100
MTB > name c1 'X' c2 'PDF'
MTB > pdf c1 c2;
SUBC> beta 1 1 .
MTB > Plot c2*c1;
SUBC>   Connect;
SUBC>   Title "PDF of Beta Distribution for particular a
and b".
```