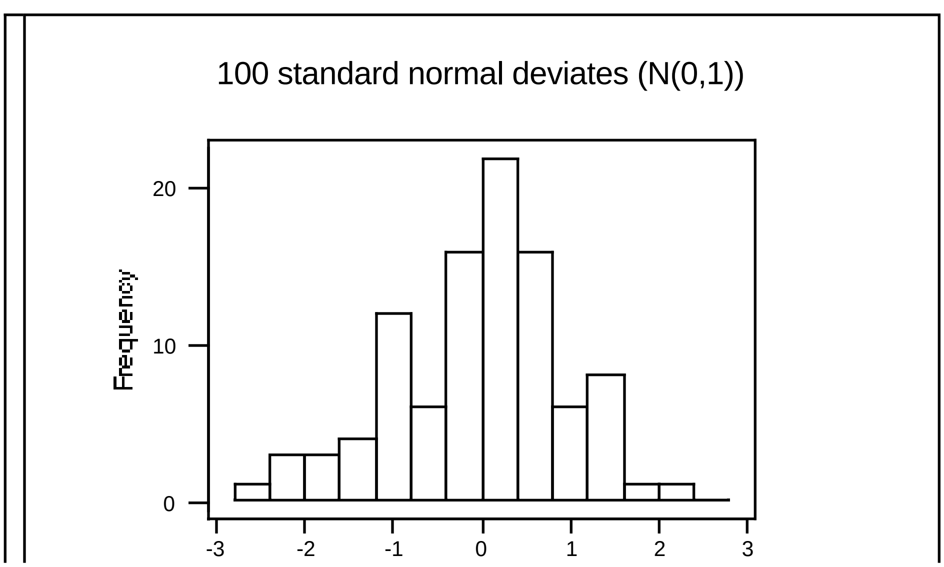
STAT 280 Nov. 28, 2001

Review of Chapters 5 & 6 today and modeling exercises.

Ch 5: Continuous Random Variables

Probability law of a continuous RV X is specified by either the pdf, $f_X(x)$, or the cdf, $F_X(x) = P(X \le x)$.
The derivative of $F_X(x)$ is $f_X(x)$.
The $f_X(x)$ values are best interpreted as relative frequencies, that is by comparing $f(x_1)$ and $f(x_2)$ by using $f(x_1)/f(x_2)$ since the probability for any particular x is 0.
For example, if $f(x) = cx^2$ on (0,3) and 0 otherwise, where c a constant, than $f(2)/f(1) = 4$ meaning that 2 is 4 times more likely than 1.
Note what is c? integrating gives c.= 1/9 so c=9

Consider Normal PDF for mean 0 SD =1
$f_X(x) = (1/\sqrt{2\pi}) \exp(-x^2/2)$
How likely are values like 0 compared to values like 1?
Look at ratio $f_x(0)/f_x(1) = 1/\exp(-1/2) = 1.65$
A histogram of normal sample values would be about 65% higher at 0 than at 1.



Where is the point of inflection of $f_X(x)$ for N(0,1)?
Second derivitive of $f_X(x)$ is 0 at –1 and +1. Helps draw.

Expected value: E(X)
Just the average of X over its population frequencies.
$(x_1 + x_2 + x_3 + \ldots + x_n)/n$

If population is 0,0,1,1,1,2,2
Then E(X) = (0+0+1+1+1+2+2)/7 = 1

Alternative method
Since there are 2 0s, 3 1s, and 2 2s we could have computed the average as (2 x 0 + 3 x 1 + 2 x 2)/7
= 0 x (2/7) + 1 x (3/7) + 2 x (2/7)
$= \sum_x xP(X = x)$
which is the formula for E(X) for discrete data.
For continuous data
$E(X) = \int_x x f_X(x)\,dx$
For example $f(x) = 8/x^3$ for $x \geq 2$, $E(x) = \int_2^\infty x(8/x^3)dx = 4$

What about $E(X^2)$?

Three methods:

i) find $f_{X^2}(x)$ via $P(X^2 \leq t)$ and differentiate wrt t, then use definition of E().

ii) Use Jacobian approach (p 207) to get $f_{X^2}(x)$ and then use definition of E()

iii) Use $E(X^2) = \int_2^{\infty} x^2 f_X(x)dx$

However in this particular example, $E(X^2)$ is infinite!
Method iii) is usually best.

Note: Var X = $E(X^2) - E^2(X)$ when these terms exist.

A note on Joint Distributions and Independence:
Consider the discrete distributions of X and Y as follows:
Suppose

X	Y	Probability
1	1	.1
1	2	.2
1	3	.1
1	4	.2
2	1	.1
2	2	.1
2	3	.1
2	4	.1

Are X and Y independent?
Is the distribution of Y|X=1 the same as the distribution of Y?
Dist of Y?

1	.2
2	.3
3	.2

4   .3

If X and Y were indept, need Y|X=1 to be the same distribution as Y|X=2. Not true here.

Same idea applies to joint continuous distributions.

Note $f_{Y|X}(y|x)$ is by definition = $f_{X,Y}(x,y)/f_X(x)$
and the definition of independence is $f_{X,Y}(x,y)= f_X(x)\, f_Y(y)$
so if X and Y are independent $f_{Y|X}(y|x) = f_Y(y)$

Ch 6: Special continuous RVs

Normal – sums and averages
Gamma – sums of exponentials
Weibull – lifetimes with non-constant hazard
Beta – distributions over a fixed interval (like (0,1))
Uniform – uniform on (a,b) e.g. U(0,1)

Gamma is approx normal for large shape parameter α
(large number of iid exponentials are added together).
Gamma with shape parameter α =1 is exponential
Weibull with shape parameter β =1 is exponential
Beta with a=1 b=1 is U(0,1)

Exercises in Modeling:

1. X= number of books checked out of library in one hour.
2. X= number of books checked out of library in first hour

3. X(t) = status of each vehicle in a fleet of taxis: working, under repair, dead.
4. X,Y = score in soccer (or hockey or …) game
5. X(t) = annual colour most in fashion
6. X(t) = stock price
7. X(t) = number of computers in assignment lab that are working.
8. X = proportion of sand in sample of construction cement
9. X = your project ????

Friday: Farouk Office Hour 12-1 AQ 10543

Monday: last lecture. More review.