Stat 280Sept 12, 2001

On to Ch 2: Discrete Random Variables (General)

RVs - variables that take numerical values according to some random mechanism

random mechanisms: coins, dice, sports contests, lifelength, stock value, daily sales, weather, insurance claims, gas consumption, airline bookings, ...

We use X to stand for the random function which generates values often denoted by x. e.g. for coin, a rv H might take successive values h,h,t,t,h,...

The nature of a rv is summarized by its probability distibution.

Probability distibutions (and rvs) can be specified by a list:

x	P(X=x)
0	.12
1	.25
2	.34
3	.29

but we often have a functional form for these probabilites, like

P(X=x) = [.9**(x-1)]*(.1) for x=1,2,3,...

These probabilites models are a great convenience, and allow us to approximate many random phenomena. All we need to do is specify, or estimate, a parameter value in a functional form like

P(X=x) = [p**(x-1)]*(1-p) for x=1,2,3,...

To use this device however, it is necessary to know what models fit which situations. We get into this in Ch 3.

Ex 2.1-4: exam with 4 questions, 3 answers each to choose from.
X is the number correct, X = 0,1,2,3,4. Assume random choice of answers.
P(X=x)=?
(2/3)**4 =.20 for x=0
4(2/3)**3 (1/3)**1 =.40 for x=1
6(2/3)**2 (1/3)**2 =.30 for x=2

4(2/3)**1 (1/3)**3 =.10 for x=3
(1/3)**4 =.01 for x=4 sums to 1.01 becasue of rounding.

Ex 2.2-4 Box 1 1,2,3,
Box 2 1,2
X number selected from box 1 and put in box 2
Y number subsequently selected from box 2 and put in box 1
P(X=1,Y=1) = 1/3*2/3 =2/9
P(X=1,Y=2) = 1/3*1/3 =1/9
P(X=1,Y=3) = 0
P(X=2,Y=1) = 1/3*1/3 =1/9
P(X=2,Y=2) = 1/3*2/3 =2/9
P(X=2,Y=3) = 0
P(X=3,Y=1) = 1/3*1/3 =1/9
P(X=3,Y=2) = 1/3*1/3 =1/9
P(X=3,Y=3) = 1/3*1/3 =1/9

is joint distribution.
P(Y=1,2,3) = 4/9, 4/9, 1/9

is marginal

P(X=1|Y=1) = P(X=1,Y=1)/P(Y=1) = 2/9 / 4/9 = 1/2

A Note on Expected Values: Just an average of population values:

If population is {1,1,1,2,2,2,2,3,3,4} then expected value is just sum
of these numbers = 21 divided by the number of numbers = 10 or 21/10=2.1

That is, if X is a randomly selected value from this population, then
E(X) = 2.1

Note that the 21 could have been derived by

1 x 3 = 3
2 x 4 = 8
3 x 2 = 6
4 x 1 = 4

and sum is 21. So E(X) = sum of x time frequency of x divided by n
= sum of xp(x)

since p(x) = frequency of x/n

E(X**2) = sum of X**2 times p(x)

E(aX +bY) = aE(X) + bE(Y) see problem 2.4-6

Var (aX +bY) = a**2 var X + b** var Y + 2ab Cov(X,Y) Section 2.8

Note correlation is the average standardized cross product.

A note on programming MINITAB:

To simulate 1000 values of a discrete distribution like

x   1,2,3,4
P(X=x).2,.3,.4,.1

Use a program like this:

```
set c1
1,2,3,4
end
set c2
.2, .3, .4, .1
end
random 1000 c3;
discrete c1 c2.
tally c3;
percents.
histogram c3.
```

This will give you the empirical distribution of the simulated data and draw a histogram.

To move the output to a word processor, just cut and paste.
To move the graph, use the EDIT/COPY GRAPH commnad and them paste.

The program can be entered sequentially, either by commands or by menus, or as a complete program.
To do the latter, you need to put the commands in a text file,

use the additional initial commands

```
MACRO
program_name
```

and add at the end of the program

```
ENDMACRO
```

Then save the whole file into the macro folder in MINITAB,
and execute the program using the commnad %program_name.

If you cannot do this because you are using a write protected disk (as in the assignment lab) then put the commands on a diskette, and provide the path to the program. The way you run it in MINITAB is then

```
%path\program_name
```

Note for Problem 2.6-4 just use MINITAB commands

```
MEAN
STDEV
```

and note that you can do componentwise arithmetic like

```
Let C2=C1*C1 or Let C2 = C1**2
```

Section 2.9: Conditional Expected Values

$E(Y) = E(E(Y|X))$

Suppose X,Y have the following joint distribution:

X	Y	P(X,Y)
1	1	.25
1	2	.20
1	3	.10
2	1	.30
2	2	.10
2	3	.05

$E(Y) = 1 * (.25 +.30) + 2* (.20 +.10) + 3 * (.10 + .05) = .55+.60+.45 = 1.6$

Now E(Y|X=1) = (1*.25 +2*.20+3*.10)/.55 = .95/.55 = 1.727
and E(Y|X=2) = (1*.30 +2* .10 +3*.05)/.45 = .65/.45 = 1.111

But E(Y)=E(E(Y|X))=(1.727*.55 + 1.444*.45) =.95+.65 = 1.6