Homework III
41000 - Business Statistics
Spring 2013
Solution

Problem I: Suppose you roll a ***fair*** die 90 times and let X be the number of times the result is greater than or equal to 5. Compute Pr(X is in [20,25]) by
a) Exact Binomial derivation (Excel function BINOMDIST);
b) Approximate normal derivation (Excel function NORMDIST).

Here, n=90 trials and p=1/3 (greater than or equal to 5 when rolling a fair die). Also, Pr(X is in [20,25]) = Pr(X=20)+Pr(X=21)+…+Pr(X=25).

Exact solution:

BINOMDIST(20,90,1/3,FALSE)+…+BINOMDIST(25,90,1/3,FALSE)

X	20	21	22	23	24	25	SUM
PR(X)	0.0069	0.0115	0.0180	0.0266	0.0372	0.0491	0.1493

Therefore, Pr(X is in [20,25])= 0.1493

Approximate solution: NORMDIST(25,m,sd,TRUE)-NORMDIST(20,m,sd,TRUE), where m = 90*(1/3)=30 and s=sqrt(90*(1/3)*(2/3))= 4.472135955.

NORMDIST(25,30,4.472135955,TRUE)= 0.1318
NORMDIST(20,30,4.472135955,TRUE)= 0.0127

Therefore, Pr(X is in [20,25]) is approximately 0.1191

c) Repeat a) and b) for Pr(X is in [35,40]).

Exact solution:

X	35	36	37	38	39	40	SUM
PR(X)	0.0467	0.0357	0.0261	0.0182	0.0121	0.0077	0.1465

Therefore, Pr(X is in [20,25])= 0.1465

Approximate solution:

NORMDIST(40,30,4.472135955,TRUE)= 0.9873
NORMDIST(35,30,4.472135955,TRUE)= 0.8682

Therefore, Pr(X is in [20,25]) is approximately 0.1191

Final comment: Notice that the exact solutions are almost the same while the approximate solutions are the same. That is because the approximate solution is based on a normal approximation centered around 30 where the intervals [25,30] and [35,40] are symmetrically located away from the mean 30.

Problem II: Ask 20 friends, colleagues and family members to answer YES or NO to each one of the following 5 questions.

Question	Sample proportion of YESES
1. Are you in favor of the death penalty?	0.566
2. Are you in favor of gun ban?	0.564
3. Do you know the capital of Argentina?	0.580
4. Do you support gay marriage?	0.720
5. Do you think statistics is useful?	0.836

For each one of the 5 questions, compute the 95% confidence interval for the true proportion of YESES (p).

I randomly selected 25 students and counted the number of yeses out of 500 trails (20 trials per homework) for each one of the above 5 questions. The 95% C.I. for p for each one of the above questions are

Lower limit = phat – 2*sqrt(phat*(1-phat)/500)
Upper limit = phat + 2*sqrt(phat*(1-phat)/500)

phat	Lower limit	Upper limit
0.566	0.522	0.610
0.564	0.520	0.608
0.580	0.536	0.624
0.720	0.680	0.760
0.836	0.803	0.869

Final remark: Some students reported yeses and noes for each question, as opposed to the final counts. That additional information allowed me to compute the following sample proportions (based on 200 trails, or 10 students' interviews):

Death penalty	Gun Ban		
	0	1	Total
0	0.165	0.260	0.425
1	0.235	0.340	0.575
Total	0.400	0.600	1.000

Pr(favor of gun ban)=0.6 (200 trials)
Pr(favor of gun ban|against death penalty)=0.26/0.425=0.612 (85 trials)
Pr(favor of gun ban|favor death penalty)=0.34/0.575=0.591 (115 trials)

These three probabilities are NOT the same, so at first glance the opinions regarding gun ban and death penalty should be considered DEPENDENT. However, what happens when we DO take into account the uncertainties when estimating 0.6, 0.612 and 0.591? Let us construct the 95% interval for the three probabilities:

n	phat	Lower limit	Upper limit
200	0.600	0.531	0.669
85	0.612	0.506	0.718
115	0.591	0.499	0.683

All intervals highly overlap with all including [0.531;0.669], so we CAN NOT confidently say these three probabilities are NOT the same. At last not based on 200, 85 and 115 trials.