p. 401 {2,5,14,32,38}

2. Do the situations involve Bernoulli trials?

Bernoulli trials if…

There are two possible outcomes.
The probability of success is constant.
The trials are independent.

a.) Rolling 5 dice, need at least two 6's to win the game. Either I get at least 2 6's or I don't, so there are two possible outcomes for each roll. The probability of getting at least 2 6's remains constant from roll to roll. Each roll is independent. Therefore, yes, these are Bernoulli trials.

b.) No, we record the distribution of eye colors. There are more than two possible eye colors, so our distribution would most likely include Blue, Green, Brown, and Hazel eyes.

c.) Dolls are defective or not. (two outcomes)
For any particular doll, the probability of being defective is (close enough to) 3%. In this context, we can assume the company made lots of dolls, say 100,000. In this case there would be 3,000 defective. So even though as defective dolls are returned, subsequent dolls have roughly a 3% chance of being defective…

$$\frac{3000}{100000} \approx \frac{2999}{99999} \approx \frac{2998}{99998} \dots$$

Therefore, as long as the dolls shipped to this local store were made independently of one another, yes these would be Bernoulli trials.

d.) No, although there are two possible outcomes here. The probability of choosing a Democrat changes (enough to matter) each time someone is selected.

e.) Students responded as cheat or not. (two outcomes)
Probability of success (cheats) is supposed to be .74 for any student.
Student responses should be independent of one another (unless they are cheaters in which case maybe they cheat on the surveys too ☺) However, I think the issue here is that the entire student body is surveyed. Therefore, I think the probability of success changes enough throughout the process to matter. Overall, no, I don't think this would be a Bernoulli trial. However, I think you could make a good argument the other way as well.

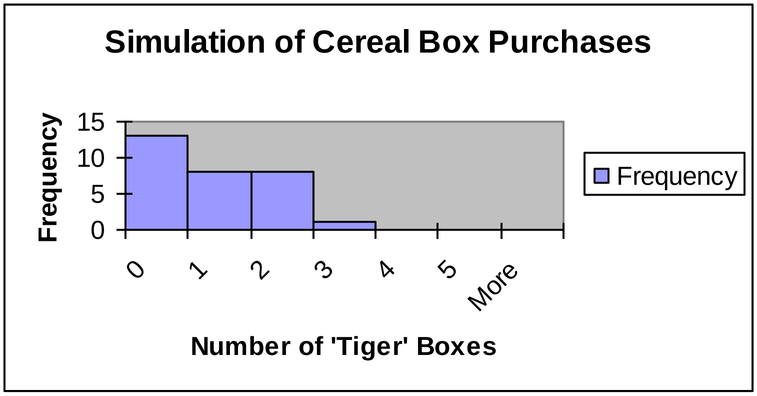
5. a) (Simulation) Generate a five random numbers 1 – 10. Let 1,2 represent a cereal box with Tiger's picture on it, while 3 – 10 represent boxes without Tiger's picture. For each string of five numbers (five boxes), we'll record the total number of 1's and 2's. Each string of numbers represents 1 trial, the total number of 1's and 2's represents the number of 'Tiger' boxes we got.

b) Using Excel, here are the first few trials and results:

Boxes	Tiger / Not	Boxes2	Tiger / Not3	Boxes4	Tiger / Not5
1	1	5	0	9	0
1	1	4	0	1	1
7	0	8	0	9	0
2	1	7	0	5	0
3	0	6	0	10	0
Number of Boxes	3	Number of Boxes	0	Number of Boxes	1

Here are the complete results of my full 30 trials.

Bin	*Frequency*	*Rel. Freq*
0	13	0.433333
1	8	0.266667
2	8	0.266667
3	1	0.033333
4	0	0
5	0	0
More	0	



c) From my 30 trials, my estimated probabilities are as follows:
P(no Tiger) = .433, P(1 Tiger) = .267, P(2 Tiger) = .267, P(3 Tigers) = .033
P(more than 4) = 0

d) To find the actual probability model, consider that this situation will have a Binomial Distribution, with probability of a Tiger box for each box equal to .2, and a sample of size 5. B(5,.2)

P(no tiger) = $(.8)^5 = .32768$

P(1 Tiger) = $5(.2)^1(.8)^4 = .4069$

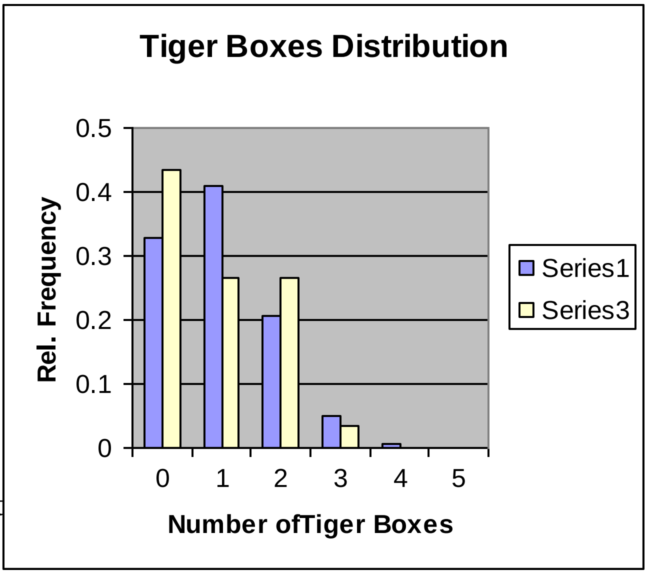
P(2 Tiger) = $\binom{5}{2}(.2)^2(.8)^3 = .2048$

P(3 Tiger) = $\binom{5}{3}(.2)^3(.8)^2 = .0512$

$$P(4\text{ Tiger}) = \binom{5}{4}(.2)^4(.8)^1 = .0064$$

$$P(5\text{ Tiger}) = (.2)^5 = .00032$$

e) Where series 1 represents the expected values and series 3 represents the values observed in my 30 trials, the distributions are pretty similar in terms of shape – they are both skewed toward the higher number of Tiger boxes. Further, they have the same median (1 Tiger box). They do, however, have different modes. The simulation returned the most 'no Tiger box' samples with 43%, while for a large number of trials, we expect the '1 Tiger box' samples to happen most often at about 41%. Although the range of Tiger boxes from my simulation is from 1 to 3 boxes, we would not consider this too different from the expected values to be cause for concern. For instance, we expect to get 4 tiger boxes about 6 out of every 1000 trials (sets of 5), so getting 0 out of 30 is not too alarming.



14. Here we have two possible outcomes (gives more than \$100, does not give more than \$100). We can assume that each person Justine contacts will elect to give (or not) independently of whether the previous person gave (or not). The probability of success (person gives more than \$100) is constant for each person. Therefore, this describes a set of Bernoulli trials, since there is not a fixed number of trials and we are interested in the number of people needed to call before Justine's first success, we have a Geometric model. We know P(gives more than \$100 given someone donates) = .05. Further, we know that P(someone donates) = .2. Thus we can find that

P(someone donates AND gives more than \$100) = (.2)(.5) = .01 = p

Therefore, E(number of calls until first \$100) =

$$\frac{1}{p} = \frac{1}{.01} = 100$$

32. There are two possible outcomes (ticket holder shows or does not). Each ticket holder (it can be assumed) shows / no shows independent of other ticket holders. This assumption is a bit shaky since it reasonable to conclude that many tickets are purchased by couples, family, or companions. However, it is probably ok to proceed. Of 275 ticket holders, the probability of success does not stay constant .05 as people no show. However, we do expect at least 10 successes (no shows) and 10 failures (people who fly) so we can use the normal distribution as an approximation to our binomial model.

$$\mu = (.05)(275) = 13.75\, noshows$$
$$\sigma = \sqrt{(275)(.05)(.95)} = 3.614$$

Therefore, a normal model with mean 13.75 no shows and standard deviation 3.614 no shows will be used.

The airline will only have enough people if at least 10 ticket holders no show. Therefore, P(enough seats) = P(no shows ≥ 10)

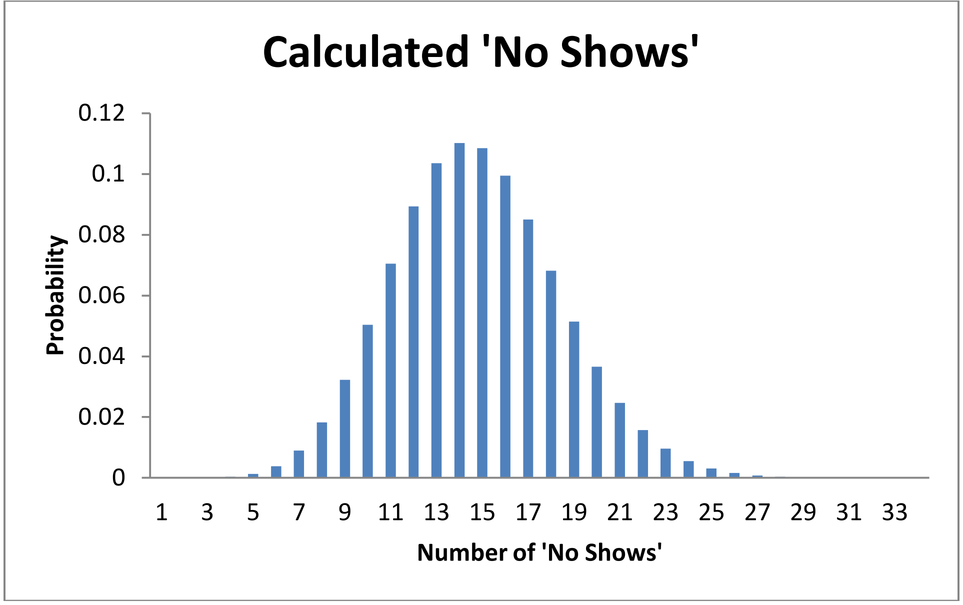
10 no shows is just beyond one standard deviation below the mean since

$$\frac{10 - 13.75}{3.614} = -1.04$$

Therefore, P(no shows $< 10) = P(z < -1.04) = .1497$

So there is about an 15% chance that the airline will ruin someone's travel plans. ☺

Here is an additional graph. It was created in Excel with actual probabilities to verify that the normal model is indeed a good approximation.



*This is an alternate method to the one chosen by students in class during the 11-12 school year, in which we used the binomcdf(…) function in the calculator to find the exact probability (about .115) it is ok that this differs by a few percent. The claim is not that the distribution of no shows IS Normally distributed, only approximately Normally distributed.

38. There are two possible outcomes for each question (True / False). Each question (if guessing) is correct or incorrect independently of the others. The probability of success on each question is constant at .5. Since we expect (if guessing) a student to get at least 25 correct and 25 incorrect, we can use a normal distribution with mean 25 questions and standard deviation of $\sqrt{(50)(.5).5)} = 3.536$ questions.

As a standard rule, we think any outcome beyond two standard deviations from the mean significant different than our expected variation. Therefore, if the student scored
25 + 2(3.536) = 32.072, so 32, questions correct, we can be (about) 97.5% confident that the student did not merely guess.

Again, our confidence level is based on the idea that the number of correctly ‘guessed’ answers is approximately Normally distributed. This assumption is valid since we expect more than 10 success (25 correct answers) and more than 10 failures (25 incorrect) answers.

Another way to view our level of confidence is this…
Only about 2.5% of students that guess on all 50 questions can expect to score as well or better than 32.