Historical dice from Asia.

Chapters 16 and 17 – Random Variables and Probability Models

Pick a Card, Any Card...

- Cost to play: $5
  - Draw the ace of hearts – Win $100
  - Draw any other ace – Win $10
  - Draw any other heart – Win $5
- Anyone willing to play?
- What if top prize is $200? $1000?

The 'Perfect' Theoretical World

- In the theoretical world, if we play the game 52 times (with the $100 top prize), what can we expect to happen for a player?

$ Outcome				
How Many Times?				

- Average those results:
- Is it even possible to win that amount?
- What does that amount mean?

Random Variables

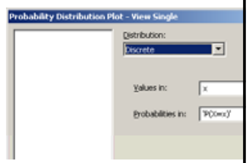
- A **random variable** assumes any of several different numeric values as a result of some random event.
  - Usually denoted with capital letters at the end of the alphabet: X, Z
  - We say the random variable is **discrete** if it can take on of a finite number of distinct outcomes
  - For the game, the amount won (lost) is the random variable.
- The **probability model** is the collection of all possible outcomes and their corresponding probabilities.

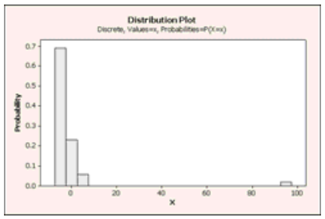
x				
P(X = x)				

Graph of a Probability Model

C1	C2
x	P(X=x)
95	0.019
5	0.058
0	0.231
-5	0.692

- Need Minitab 15 (or sketch by hand)
  - Put the x's in a column
  - Put corresponding P(X = x) in another
  - Graph > Probability Distribution Plot > View Single
    - Distribution: Discrete
    - Values in: Column with x's
    - Probabilities in: Column with probabilities





- Where is the mean (center of gravity)?

Expected Value

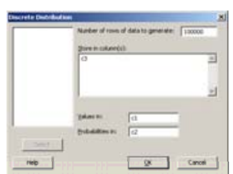
- The **expected value** of a random variable is the theoretical long-run average.
  - Denoted by E(X) or μ.
  - E(X) gives us a way of measuring the *center* of the probability model.
  - Calculated by summing the products of the variable values and probabilities.
- Random variables have **standard deviations** and **variances** too
  - Denoted SD(X) and Var(X), respectively
- Calculating exact values of E(X) and SD(X) is tedious in Minitab, so we will rely on the LLN to *estimate* them.

Example: Calculating E(X) by hand

C1	C2
x	P(X=x)
95	0.019
5	0.058
0	0.231
-5	0.692

Estimating E(X) and SD(X) with Minitab

- Enter the outcomes (X) in C1
- Enter the probabilities (P(X=x)) in C2
- Generate a LARGE amount of samples from the probability distribution
  - Calc > Random Data > Discrete
  - Number of Rows: 1000000
  - Store in: C3
  - Values in: C1
  - Probabilities in: C2



- Then calculate the mean and SD of your freshly sampled list
  - Use Stat > Basic > Display Descriptive Statistics

Just Checking…

- Every time you repeat this, your answers will change, but they should all be close to the theoretical values…

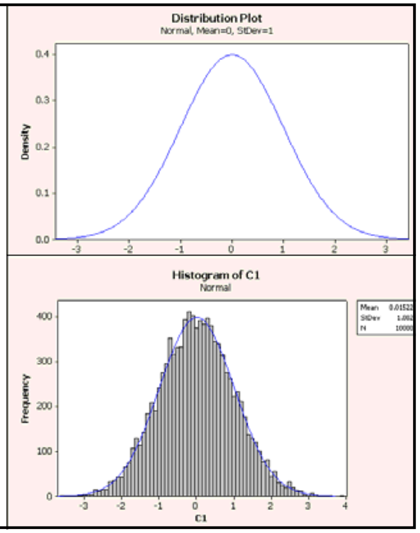
```
Descriptive Statistics: Samples

Variable     Mean    StDev  Variance
Samples   -1.3636  13.7233  188.3301
```

- Note: The true theoretical value for E(X) was -1.35
  - Our answer was off for two reasons: (1) we sampled and (2) we used rounded probabilities
- More about spread…
  - The spread gives a measure for how far values are away from the center.
  - Would you have preferred a small SD or a large one in the card game?

LLN in Action N(0,1)

- The LLN guarantees that with enough samples, the simulated distribution will be 'close enough' to the real distribution.
- Top: N(0,1)
- Bottom: 10,000 samples from N(0,1)



Special Distribution: Bernoulli Trials

- Bernoulli trial requirements
  - Only two possible outcomes: Success or Failure
    - Yes or No
  - Probability of success is the same for each trial (constant), p
  - Trials are independent
- Abbreviated: Bernoulli(p)
- Some Examples – Verify the requirements
  - Trial: Flip a coin. Success: Heads. Failure: Not Heads.
  - Trial: Roll a die. Success: Roll a 6. Failure: Anything but a 6.
  - Trial: Roll a die. Success: Even number. Failure: __________.
  - Trial: Ask registered voter if they voted for the current president. Success: _____. Failure: ___________.

Mean and SD of Bernoulli Trials

- We don't need to estimate the mean/SD for Bernoulli trials…
- Mean of Bernoulli Trials = E(X) = p
- Standard Deviation = $\sqrt{p(1-p)}$
  - Often (1-p) is written as q (to specify the probability of failure), so the standard deviation can be written as $\sqrt{pq}$

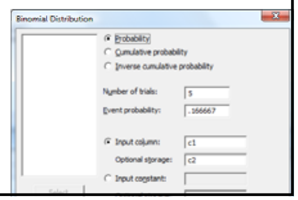
Example	p	q	Mean	SD
Trial: Flip a coin. Heads = Success				
Trial: Roll a die. A 6 = Success				
Trial: Roll a die. An even number = Success				
Trial: Ask registered voter if they voted for the current president. Obama = Success				

Special Distribution: Binomial Model

- The **binomial probability model** describes the number of successes in a specified number of Bernoulli trials.
  - Abbreviated: Binom(n,p)
  - Mean = np
  - SD = sqrt(npq)
- **Ex**: Rolling dice! We throw five dice, and we want to keep track of the number of 1's.
  - What are the possible outcomes?
  - What model describes the outcomes?
  - What is the mean of the model?
  - What is the standard deviation of the model?

Calculating Probability: Binomial Model

- **Ex**: Rolling dice! We throw five dice, and we want to keep track of the number of 1's.
- We can use Minitab to write down the probability of k successes for k = 0 to n
  - First make a list of all the possible outcomes
    - Calc > Make Patterned Data > Simple Set of Numbers
    - Store in C1
    - From 0 to n (n = 5 in this example)
  - Then find the probabilities
    - Calc > Probability Distributions > Binomial
    - Probability
    - Number of trials: n (n = 5 in this example)
    - Event probability: 1/6 ~ .166667
    - Input Column: C1, Storage: C2
    - Note: We could have entered each outcome one at a time in "Input constant"
  - Repeat the last step, storing Cumulative Probabilities in C3



Just Checking…

- So… What does "cumulative probability" mean?

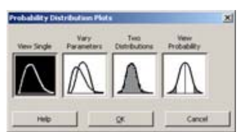
Worksheet 1 ***

↓	C1	C2	C3
	Outcome	Prob	CumulProb
1	0	0.401877	0.40188
2	1	0.401878	0.80375
3	2	0.160751	0.96451
4	3	0.032150	0.99666
5	4	0.003215	0.99987
6	5	0.000129	1.00000

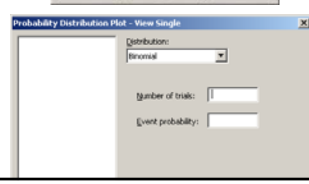
- **Ex**: Rolling dice! We throw five dice, and we want to keep track of the number of 1's.
  1. What is the chance of getting five 1's?
  2. What is the chance of getting five or fewer 1's?
  3. What is chance of getting at least one 1?

Graphing the Binomial Model

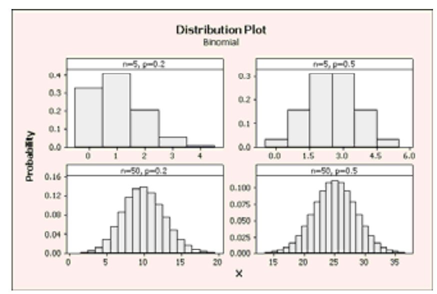
- With Minitab 15, we can graph Binom(n,p) for different values of n and p
- Minitab > Graph > Probability Distribution Plot
  - View Single
  - Distribution: Binomial
  - Number of Trials: n
  - Event Probability: p



- Graph the following:
  1. Binom(5, .2)
  2. Binom(50, .2)
  3. Binom(5, .5)
  4. Binom(50,.5)



Just Checking…



- *Note: You can use Probability Distribution Plot > View Probability to calculate probabilities for the Binomial model like we did for the Normal*
  - *Much easier than what we did, but only works on Minitab 15*

Normal Approximation to Binomial

- So... when can we use the normal approximation?
  - We must expect to have at least 10 successes and 10 failures
  - ie: $np \geq 10$ and $nq \geq 10$
- And what normal is it?
  - For np and nq large enough, Binom(n,p) ~ Normal(np, $\sqrt{npq}$)
- Binom(50, .5) ~ ____________
- Binom(50, .2) ~ ____________

Class Work

- To get credit, it is your responsibility to get checked off.
  1. Chapter 14 Handout
     - Checking solutions? No pens in the front!
  2. There is class time to work on your project after class work…

Homework

- Textbook/Routine Homework
  - Due Next Week (25% chance of collection)
  1. Read Chapter 16 and 17
  2. Pg 427-431: #3, 5, 17, 21
  3. Pg 447-450: #1, 17, 25, 27, 28, 29, 31
  4. Gradekeeper Assignment

Answers to Even #s
C17 #28
(a) 4, (b) 3.2, (c) 0.894

- Project/Exploration Homework
  - Project #2 – Surveys Due Next Time!!!

Exam 2 in 1 Week!!!

- Review Next Time!