Stat 1040
Review 2

1. Three draws are made at random from the box [3, 4, 4, 5, 5, 5].

a) If the draws are made with replacement, find the probability that a "4" is drawn each time.

$\frac{2}{6} \cdot \frac{2}{6} \cdot \frac{2}{6} = \frac{1}{27} = \frac{1}{27}$

b) If the draws are made without replacement, find the probability that a "5" is drawn each time.

$\frac{3}{6} \cdot \frac{2}{5} \cdot \frac{1}{4} = \frac{1}{20}$

c) If the draws are made without replacement, find the probability that a "3" is drawn at least once.

$1 - P(\text{no "3s"}) = 1 - \frac{5}{6} \cdot \frac{4}{5} \cdot \frac{3}{4} = \frac{1}{2}$

2. Suppose that 60% of all people who are eligible for jury duty in a large city are in favor of capital punishment. We are interested in how this fact might affect the composition of a jury in a murder trial. Suppose a jury of 12 is to be randomly selected from all who are eligible for jury duty in that city.

a) What is the chance that none of the 12 jurors selected favors capital punishment?

$(.4)^{12} = .0000167$ .0022.

b) If the jury were selected at random, would you be surprised if none of the 12 jurors selected favored capital punishment? Explain.

Yes! ≈ 17 in a million

3. Shaquille O'Neal has a lifetime 53% chance of making a free throw. He says he can't shoot free throws because he fell out of a tree as a child and broke both his wrists. You may assume that this chance is constant and that all of his free-throws are independent of each other. Shaq takes 10 free- throws.

a) What is the chance that he makes none of the 10 shots?

$(.47)^{10} = .000526$

b) What is the chance that he makes exactly 7 of the 10 shots?

Repeated trials $n = 10$ $k = 7$ $p = .53$

$\binom{10}{7}(.53)^7(.47)^3 = \frac{10 \cdot 9 \cdot 8}{3 \cdot 2}(.53)^7(.43)^3$

$= .142$ 14.2%

4. In one play of a game, a nickel, dime, and quarter are tossed. If all three come up "heads", then you win \$ 15 ; otherwise, you lose \$ 3 . You repeat this game 100 times.

a) What is the probability of getting "three heads" in one play of the game?

$\frac{1}{8}$

b) Find an appropriate box model for determining your total winnings after 100 plays.

$\boxed{\$15,\ 7\ -\$3s}$ → Draw 100 + consider the sum of draws.

$\text{Box SD} = \sqrt{\frac{1}{8} \cdot \frac{7}{8}}\ (15 - (-3))$

Note:
EV = − \$75
SE = \$ 59.53

5. Find the chance of getting four aces in a hand of five cards.

AAAAN
AAANA
AANAA
ANAAA
NAAAA

$\frac{4}{52} \cdot \frac{3}{51} \cdot \frac{2}{50} \cdot \frac{1}{49} \cdot \frac{48}{48}$

$\frac{48}{52} \cdot \frac{4}{51} \cdot \frac{3}{50} \cdot \frac{2}{49} \cdot \frac{1}{48}$

$5\left(\frac{1}{13} \cdot \frac{1}{17} \cdot \frac{1}{25} \cdot \frac{1}{49}\right)$

6. A gambler is going to play roulette 25 times, putting a dollar on a *split* each time, say *11 or 12*. If either number comes up, the gambler gets the dollar back, together with winnings of \$17, otherwise he loses the dollar. So a *split* pays 17 to 1 . Build a box model for the gambler's net gain.

$\boxed{2\ \$17,\ 36\ -\$1s}$ → Draw 25 times + consider the sum of draws

or $\boxed{\$17,\ 18\ -\$1s}$

7. The Heart Association claims that only 10% of adults over 30 years of age in the country can pass the minimum fitness requirements established by the President's Physical Fitness Commission. Suppose that 100 adults over 30 years of age are randomly selected and all are given the fitness test. Assuming the Heart Association's claim is true, find the probability that at least 16% of them will pass the minimum fitness requirements.

$\boxed{10\%\ 1s,\ 90\%\ 0s}$ → Draw 100 + consider the % 1s drawn.

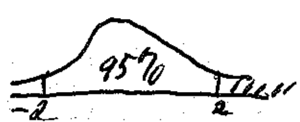
$\text{Box SD} = \sqrt{(.1)(.9)} = \frac{3}{10}$

EV for % = 10%

$\text{SE for \%} = \frac{\text{Box SD} \times \sqrt{100}}{100} \times 100\%$

= 3%

$\text{Normal} \approx \quad \frac{16-10}{3} = 2$,



$2\frac{1}{2}\%$

8. A box contains five tickets: two 2's, two 3's, and one 4. We draw from this box with replacement. Fill in the blanks and explain:

a) As the number of draws gets larger and larger, the data histogram of the draws will look more and more like <u>the histogram for the box</u>.

b) As the number of draws gets larger and larger, the probability histogram for the <u>average</u> of the draws (when put in standard units) will look more and more like <u>standard normal curve</u>

9. Draw 100 times with replacement from the box [1 , 3 , 7 , 13].

Box AV = 6

Box SD = $\sqrt{21} = 4.58$

EV for sum = 600

SE for sum $= (4.58)\sqrt{n}$

$= 45.8$

a) How small can the sum of the draws be? How large?

100 , 1300

b) How many times do you expect to draw a "7"?

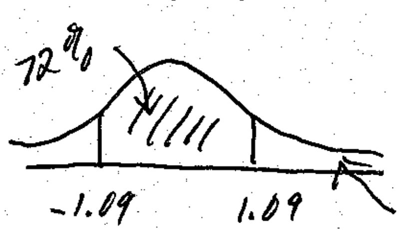
$\frac{1}{4}$ th of the time <u>25 times</u>

c) What do you expect the sum of the draws to be?

Box AV × 100 = 600

d) Find the probability that the sum of the draws is larger than 650.

Normal ≈ $\dfrac{650-600}{(4.58)10} \approx 1.09$



$A(1.09) = 72\%$

$\frac{28}{2} = 14\%$

10. Over 200,000 Utah adults are employed part-time and about 4 times as many are employed full-time. To estimate the percentage of citizens who lack health insurance, the state government wants to take a simple random sample. In each case, if all other things are similar, which is more accurate or are they about the same?

i. A simple random sample of 1000 part-time employed adults.
ii. A simple random sample of 1000 full-time employed adults. About same

i. A simple random sample of 1000 part-time employed adults.
ii. A simple random sample of 4000 full-time employed adults. ← More accurate

i. A simple random sample of 2% of the population of the part-time employed.
ii. A simple random sample of 2% of the population of the full-time employed. ← More accurate

$\frac{346}{500} \times 100\% \approx 69\%$

11. In a simple random sample of 500 Cache Valley drivers, 346 claim that they always wear a seatbelt. Find a 95% confidence interval for the percentage of *all* Cache Valley drivers who always wear a seatbelt.

$\boxed{?\ 1s,\ ?\ 0s} \rightarrow$ Draw 500 + consider % 1s

Box SD ≈ Sample SD $= \sqrt{\frac{346}{500} \cdot \frac{154}{500}} = .46$

EV for % 1s = Box %

SE for % 1s = $\frac{\text{Box SD} \times \sqrt{500}}{500} \times 100\% = 2\%$

69% ± 2 SE

69% ± 4%

(65, 73)

12. Lake Taupo is a New Zealand lake with only two types of fish; brown trout and rainbow trout. A wildlife expert takes a simple random sample of 250 fish from the lake and finds that 47 of them are brown trout. Find a 90% confidence interval for the percentage of brown trout in the lake. (You may assume the number of fish in the lake is very large.)

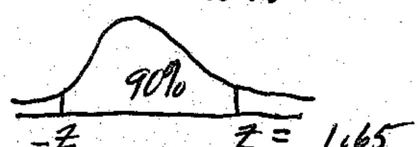
$\frac{47}{250} \times 100\% = 18.8\%$

Box SD $\approx \sqrt{\frac{47}{250} \cdot \frac{203}{250}} \approx .39$

SE for % 1s $= \frac{(.39)\sqrt{250}}{250} \times 100\% = 2.47\%$

90% C.I

$A(z) = 90\%$



18.8% ± (1.65)(2.47)%

18.8% ± 4%

13. Sixty four randomly selected high school seniors from the school district are given the math portion of the ACT test with the following results: AV = 22 , SD = 4 .

a) Construct the 95% confidence interval for the average ACT score of all seniors in the school district.

$\boxed{\text{ACT scores}} \rightarrow$ Draw 64 + consider the AV of draws

Box SD ≈ Sample SD = 4

EV for AV = Box AV

SE for AV $= \frac{\text{Box SD} \cdot \sqrt{64}}{64} = \boxed{.5}$

22 ± 2 SEs , 22 ± 2(.5)

$\boxed{22 \pm 1}$ $\boxed{(21, 23)}$

b) If you wanted a 95% confidence interval with half the length of the interval in part a), you would <u>Sample 4 times as many , 256</u>

c) About 95% of the seniors have an ~~IQ~~ ACT in the interval <u>22 ± 2(4)</u>.

14. The women's center at a local hospital has a record of all deliveries in the last 10 years (about 12,000). They take a simple random sample of 500 of these deliveries and find that the average gestation time is 266 days with an SD of 16 days.

a) Find an approximate 95% confidence interval for the average gestation time for all 12,000 deliveries in the last 10 years.

| gestation days | → Draw 500 + consider the average of draws

Box AV ≈ Sample AV = 266

EV for AV = Box AV or Population AV

$$SE \text{ for } AV = \frac{\text{Box SD} \times \sqrt{500}}{500} = .715$$

95% C.I = 266 ± 2SE

$266 \pm 2(.715)$ $\boxed{266 \pm 1.43}$

b) The gestation time is not normal; it has a long left tail and cuts off quite sharply on the right. Does this mean that your confidence interval is incorrect? Would it be valid to use the normal curve to figure out what percentage of the gestation times were longer than 282 days?

C.I is correct; we have a quite accurate estimate of the true average gestation time.

The normal ≈ does not apply to the contents of this box.

Repeated Trials: Suppose we have n independent trials, and the probability that event E occurs in an given trial is p. Then the probability that E will occur exactly k times is

$$\frac{n!}{k!(n-k)!} p^k (1-p)^{n-k}$$

Expected Value and Standard Error

Suppose you randomly draw n times with replacement from a box.

EV for the sum of the draws = Box AV x n

SE for the sum of the draws = Box SD x $\sqrt{n}$

EV for the average of the draws = Box AV

SE for the average of the draws = $\dfrac{\text{Box SD} \times \sqrt{n}}{n}$

EV for the % of 1s drawn = % of 1s in the box

SE for the % of 1s drawn = $\dfrac{\text{Box SD} \times \sqrt{n}}{n} \times 100\%$

Short-Cut Formulas for SDs

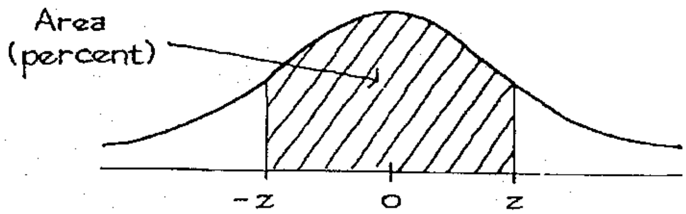
L -S Boxes: Suppose the box contains only two different numbers, Ls and Ss with $L > S$, and the proportion of Ls is equal to p and the proportion of Ss is equal to $1-p$.

Then the Box SD $=(L-S)\sqrt{p \times (1-p)}$.

Correction Factor for SEs

$$\textit{correction factor} = \sqrt{\frac{\textit{number of tickets in box} - \textit{number of draws}}{\textit{number of tickets in box} - \textit{one}}}$$

A NORMAL TABLE



z	Area	z	Area	z	Area
0.00	0	1.50	86.64	3.00	99.730
0.05	3.99	1.55	87.89	3.05	99.771
0.10	7.97	1.60	89.04	3.10	99.806
0.15	11.92	1.65	90.11	3.15	99.837
0.20	15.85	1.70	91.09	3.20	99.863
0.25	19.74	1.75	91.99	3.25	99.885
0.30	23.58	1.80	92.81	3.30	99.903
0.35	27.37	1.85	93.57	3.35	99.919
0.40	31.08	1.90	94.26	3.40	99.933
0.45	34.73	1.95	94.88	3.45	99.944
0.50	38.29	2.00	95.45	3.50	99.953
0.55	41.77	2.05	95.96	3.55	99.961
0.60	45.15	2.10	96.43	3.60	99.968
0.65	48.43	2.15	96.84	3.65	99.974
0.70	51.61	2.20	97.22	3.70	99.978
0.75	54.67	2.25	97.56	3.75	99.982
0.80	57.63	2.30	97.86	3.80	99.986
0.85	60.47	2.35	98.12	3.85	99.988
0.90	63.19	2.40	98.36	3.90	99.990
0.95	65.79	2.45	98.57	3.95	99.992
1.00	68.27	2.50	98.76	4.00	99.9937
1.05	70.63	2.55	98.92	4.05	99.9949
1.10	72.87	2.60	99.07	4.10	99.9959
1.15	74.99	2.65	99.20	4.15	99.9967
1.20	76.99	2.70	99.31	4.20	99.9973
1.25	78.87	2.75	99.40	4.25	99.9979
1.30	80.64	2.80	99.49	4.30	99.9983
1.35	82.30	2.85	99.56	4.35	99.9986
1.40	83.85	2.90	99.63	4.40	99.9989
1.45	85.29	2.95	99.68	4.45	99.9991

PROBABILITY RULES (The World)

Definition: The probability [chance] of event A is the proportion [percentage] of the time A is expected to happen when the random process is repeated over and over again.

Opposite Event Rule: The probability that event A happens is equal to one minus the probability that A doesn't happen.

Multiplication Rule: The probability that events A and B both happen is equal to the probability that A happens times the probability that B happens given that event A has occurred.

Definition: Two events are mutually exclusive when the occurrence of one prevents the occurrence of the other.

Addition Rule: The probability that event A or event B happens is equal to the probability that A happens plus the probability that B happens minus the probability that both happen. If events A and B are mutually exclusive, then the probability that event A or B happens is simply the sum of the probabilities.

Definition: Two events are independent if when one happens, the probability that the other happens is unchanged.

Fundamental Counting Principle: If event A can occur in m ways and after A occurs event B can occur in n ways, then the number of ways both events A and B can occur is m x n .

The number of ways k objects can be selected from n objects without regard to order is

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}$$

Repeated Trials: Suppose we have n independent trials, and the probability that event E occurs in an given trial is p. Then the probability that E will occur exactly k times is

$$\frac{n!}{k!(n-k)!} p^k (1-p)^{n-k}$$