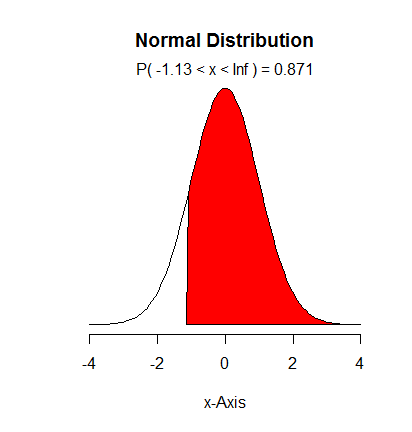
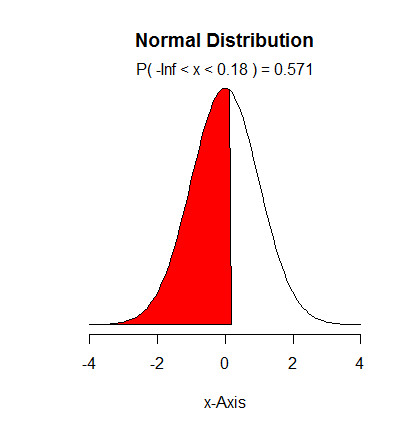
IS606 Chapter 3 HW

Andrew Goldberg

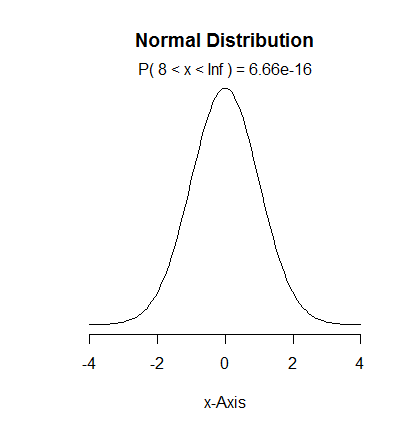
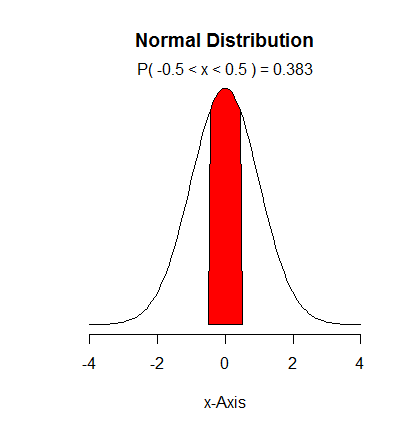
**3.2 Area under the curve, Part II.** What percent of a standard normal distribution N(μ = 0, μ = 1) is found in each region? Be sure to draw a graph.

a. Z > -1.13 = neg z = .1292 so 1-.1292 = **.871** b. Z < 0.18 = **.571**



c. Z > 8 = **very small, well under 1-in-1 million** d. |Z| < 0.5 = Z > -.5 =.3085

Z < .5 = .6915

 .6915 - .3085 = **.383**

**3.4 Triathlon times, Part I.** In triathlons, it is common for racers to be placed into age and gender groups. Friends Leo and Mary both completed the Hermosa Beach Triathlon, where Leo competed in the Men, Ages 30 - 34 group while Mary competed in the Women, Ages 25 - 29 group. Leo completed the race in 1:22:28 (4948 seconds), while Mary completed the race in 1:31:53 (5513 seconds). Obviously Leo finished faster, but they are curious about how they did

within their respective groups. Can you help them? Here is some information on the performance of their groups:

• The finishing times of the Men, Ages 30 - 34 group has a mean of 4313 seconds with a standard deviation of 583 seconds.

• The finishing times of the Women, Ages 25 - 29 group has a mean of 5261 seconds with a standard deviation of 807 seconds.

• The distributions of finishing times for both groups are approximately Normal.

**(a)** Write down the short-hand for these two normal distributions.

Men 30 - 34: **N(μ = 4131, σ = 583)**

Women 25-29: **N(μ = 5261,** σ **= 807)**

**(b)** What are the Z-scores for Leo's and Mary's finishing times? What do these Z-scores tell you?

Leo: Z = (4948 – 4131)/583 = **1.401 (can argue its -1.401 as larger scores actually mean worse)**

Leo was nearly a standard deviation and a half slower than the average runner in his field.

Mary: Z = (5513 – 5261)/807 = **.312 (can argue for -.312)**

Mary was less than a third of a standard deviation slower than the average runner in her field.

**(c)** Did Leo or Mary rank better in their respective groups? Explain your reasoning.

**Mary ranked better as she was only about a third slower than the average runner, while Leo was well over a standard deviation slower.**

**(d)** What percent of the triathletes did Leo finish faster than in his group?

Leo’s Z was (-)1.401, = **.0808** or he finished faster than **8.1%** of his group

**(e)** What percent of the triathletes did Mary finish faster than in her group?

Mary’s Z was (-).312, = **.3783** or faster than **37.8%** of her group

**(f)** If the distributions of finishing times are not nearly normal, would your answers to parts

(b) - (e) change? Explain your reasoning.

**The Z scores would remain the same, but we wouldn’t be able to answer b to e since we can’t use a normal probability table to calculate probabilities and percentiles without a normal model**

**3.18 Heights of female college students.** Below are heights of 25 female college students.

**(a)** The mean height is 61.52 inches with a standard deviation of 4.58 inches. Use this information to determine if the heights approximately follow the 68-95-99.7% Rule.

**Yes, they approximate the 68-95-99.7% rule:**

@+1std = 61.52 + 4.58 = 66.1, between student 21 and 22

@-1std = 61.52 – 4.58 = 56.94, at student 5

So 21.5 – 5 = 16

16.5/25 = **66%** within first deviation, approximating 68%

@+2std = 70.68, 24th student

@-2std = 52.36, under 1st student

So 24/25 = **96%** within second deviation, approximating 95%

@+3std = 75.26

@-3std = 47.48

This covers **100%**, but still approximates 99.7

**(b)** Do these data appear to follow a normal distribution? Explain your reasoning using the graphs provided below.

**The sample size is very small, but the histogram appears to be unimodal and approximates the histogram, and there’s only a few outliers on the probability plots, so it likely follows the normal distribution.**

**3.22 Defective rate.** A machine that produces a special type of transistor (a component of

computers) has a 2% defective rate. The production is considered a random process where each transistor is independent of the others.

**(a)** What is the probability that the 10th transistor produced is the first with a defect?

.98^9 \* .02 = **.017**

**(b)** What is the probability that the machine produces no defective transistors in a batch of 100?

.98^100 = **.133**

**(c)** On average, how many transistors would you expect to be produced before the first with a

defect? What is the standard deviation?

1/.02 = **μ = 50**

Sqrt(1-.02)/(.02^2) = **σ = 49.5**

**(d)** Another machine that also produces transistors has a 5% defective rate where each transistor is produced independent of the others. On average how many transistors would you expect to be produced with this machine before the first with a defect? What is the standard deviation?

1/.05 = **μ = 20**

Sqrt(1-.05)/(.05^2) = **σ = 19.5**

**(e)** Based on your answers to parts (c) and (d), how does increasing the probability of an event affect the mean and standard deviation of the wait time until success?

**There is a lower mean wait time with higher probability events because they happen more often, and the standard deviations are lower as well.**

**3.38 Male children.** While it is often assumed that the probabilities of having a boy or a girl are the same, the actual probability of having a boy is slightly higher at 0.51. Suppose a couple plans to have 3 kids.

**(a)** Use the binomial model to calculate the probability that two of them will be boys.

3 choose 2 = 3!/((2!)(3-2!)) = 3

(.51^2)(.49) = .128

.128 \* 3 = **.38**

**(b)** Write out all possible orderings of 3 children, 2 of whom are boys. Use these scenarios to

calculate the same probability from part (a) but using the addition rule for disjoint outcomes.

Confirm that your answers from parts (a) and (b) match.

BBG = .51 \* .51 \* .49 = .128

BGB = .51 \* .49 \* .51 = .128

GBB = .49 \* .51 \* .51 = .128

.128 \* 3 = **.38**

**(c)** If we wanted to calculate the probability that a couple who plans to have 8 kids will have

3 boys, briefly describe why the approach from part (b) would be more tedious than the

approach from part (a).

**Writing out each of the disjoint permutations of 8 total children would be very tedious, while binomial distributions are much quicker.**

**3.42 Serving in volleyball.** A not-so-skilled volleyball player has a 15% chance of making the serve, which involves hitting the ball so it passes over the net on a trajectory such that it will land in the opposing team's court. Suppose that her serves are independent of each other.

**(a)** What is the probability that on the 10th try she will make her 3rd successful serve?

9 choose 2 = 36

(.15^3)(.85^7) = .001082

36 \* .001082 = **.039**

**(b)** Suppose she has made two successful serves in nine attempts. What is the probability that

her 10th serve will be successful?

**15%, since each serve is independent.**

**(c)** Even though parts (a) and (b) discuss the same scenario, the probabilities you calculated

should be different. Can you explain the reason for this discrepancy?

**In a, we are calculating the possibilities all of the possible permutations that lead up to her serving successfully on her 10th try. In b, we are only looking at the probability of her being successful in one serve, since it is supposedly independent from her past actions.**