http://www.stat.ucla.edu/~dinov/courses_students.dir/10/Winter/STAT13.1.dir/

STAT 13 Homework 4 Solutions

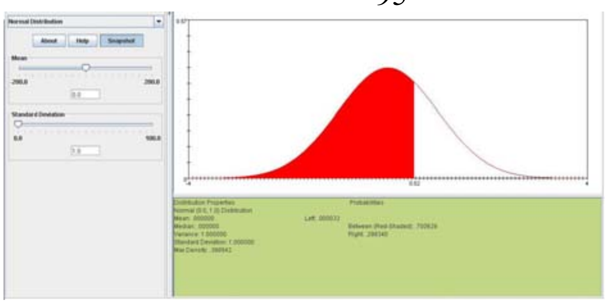
Do the following problems using the interactive SOCR Normal Distribution calculator (http://socr.ucla.edu/htmls/dist/Normal_Distribution.html). Include snapshots of your work to support your findings.

Question 1. The brain weights of adult Swedish males are approximately normally distributed with mean μ = 1,350g and standard deviation = 95g. Let Y denote the brain weight of a randomly chosen person from this population. Calculate:

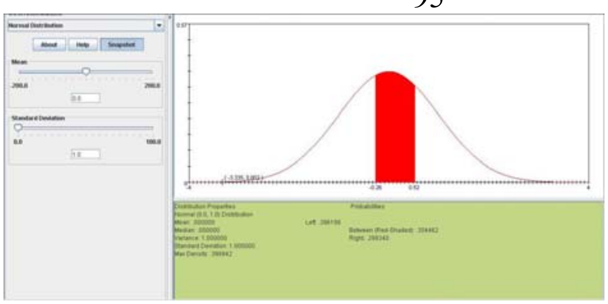
P(Y <=1,400).
P(1,325 <= Y <=1,400).
P(1,325 <= Y).
P(1,450 <= Y).
P(1,450 <= Y <=1,500).
P(1,300 <=Y <=1,350).

Answer:

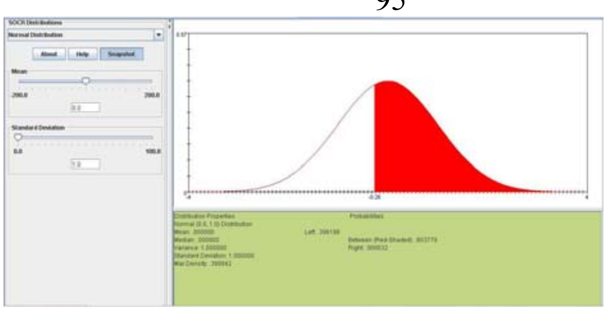
1. $P(Y \le 1400) = P(Z \le \frac{1400-1350}{95}) = P(Z \le 0.5263) = 0.7007$



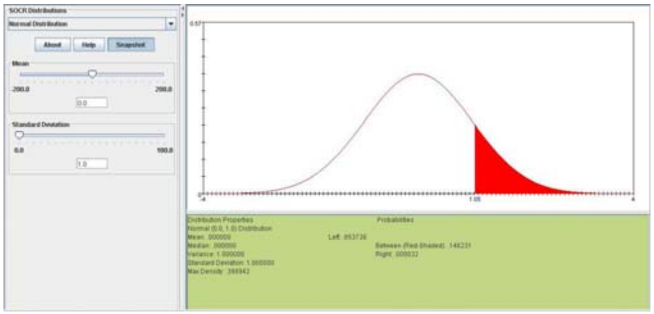
2. $P(1325 \le Y \le 1400) = P(\frac{1325-1350}{95} \le Z \le \frac{1400-1350}{95}) = P(-0.2632 \le Z \le 0.5263) = 0.3045$



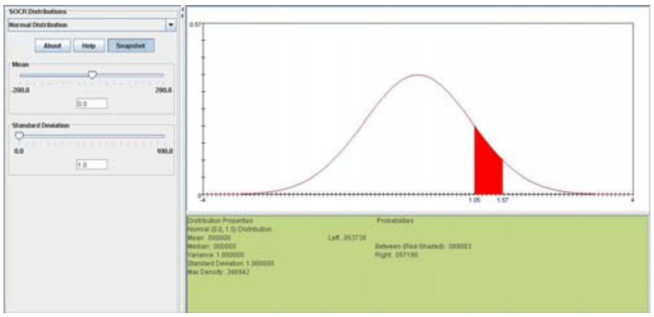
3. $P(Y \ge 1325) = P(Z \ge \frac{1325-1350}{95}) = P(Z \ge -0.2632) = 0.6038$



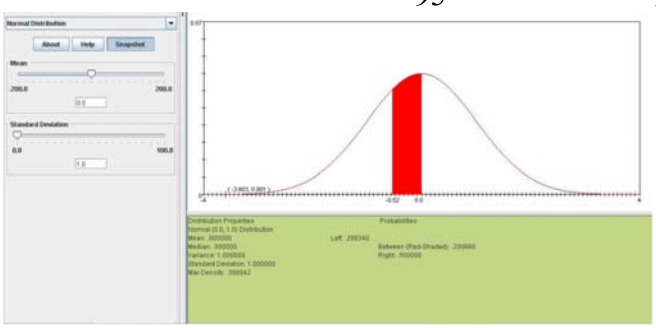
4. $P(Y \ge 1450) = P(Z \ge \frac{1450-1350}{95}) = P(Z \ge 1.0526) = 0.1463$



5. $P(1450 \le Y \le 1500) = P(\frac{1450-1350}{95} \le Z \le \frac{1500-1350}{95}) = P(1.0526 \le Z \le 1.5789) = 0.0891$



6. $P(1300 \le Y \le 1350) = P(\frac{1300-1350}{95} \le Z \le \frac{1350-1350}{95}) = P(-0.5263 \le Z \le 0) = 0.2007$



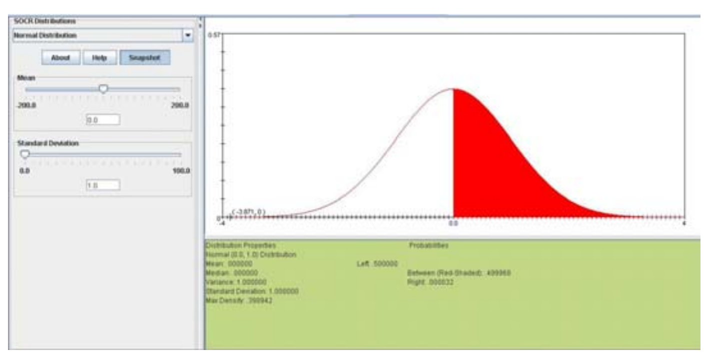
Question 2. The serum cholesterol levels of 17-year-olds follows a normal distribution with mean 176 mg/dLi and standard deviation 30 mg/dLi. What percentage of 17-year-olds have serum cholesterol values:

176 or more?
166 or less?
212 or less?
123 or more?
between 176 and 206?
between 123 and 157?
between 155 and 186?

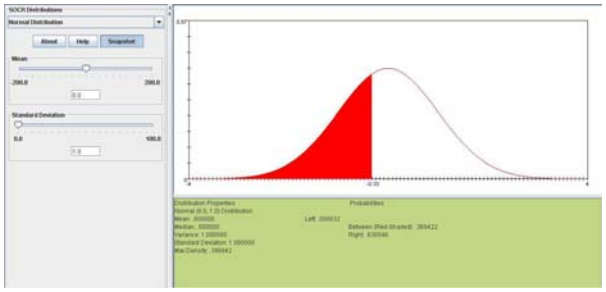
Answer:

Let X be the serum cholesterol levels of 17-year-olds. Then

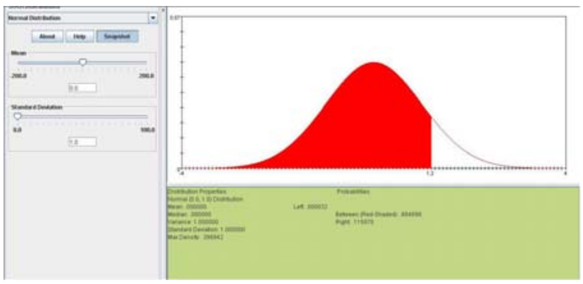
1. $P(X \ge 176) = P(Z \ge \frac{176-176}{30}) = P(Z \ge 0) = 0.5$



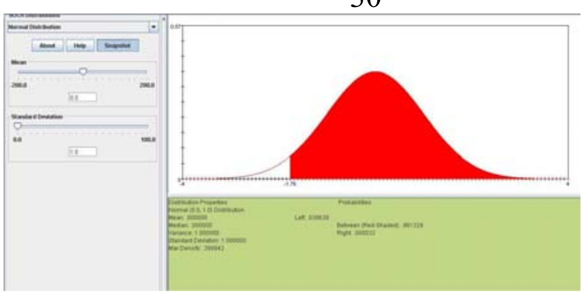
2. $P(X \le 166) = P(Z \le \frac{166-176}{30}) = P(Z \le -0.3333) = 0.3695$



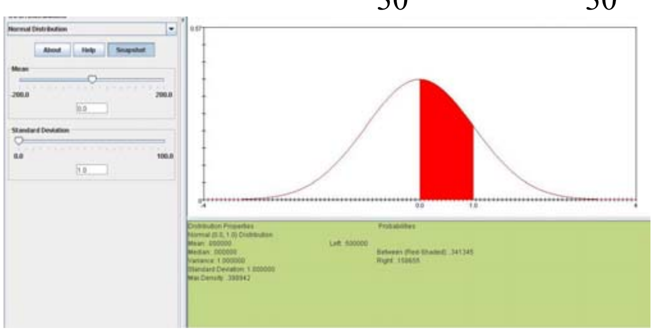
3. $P(X \le 212) = P(Z \le \frac{212-176}{30}) = P(Z \le 1.2) = 0.8849$



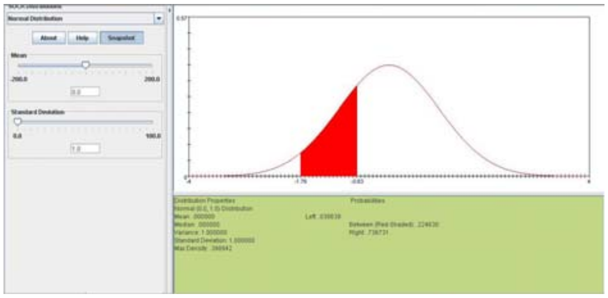
4. $P(X \ge 123) = P(Z \ge \frac{123-176}{30}) = P(Z \ge -1.7667) = 0.9614$



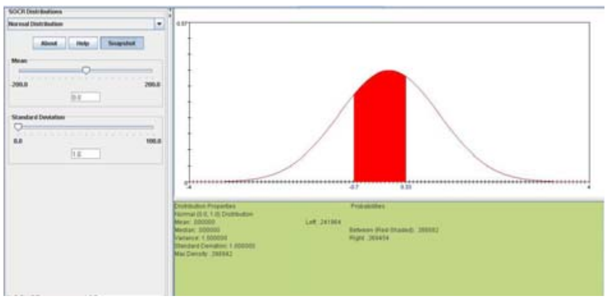
5. $P(176 \le X \le 206) = P(\frac{176-176}{30} \le Z \le \frac{206-176}{30}) = P(0 \le Z \le 1) = 0.3413$



6. $P(123 \le X \le 157) = P(\frac{123-176}{30} \le Z \le \frac{157-176}{30}) = P(-1.7667 \le Z \le -0.6333) = 0.2246$

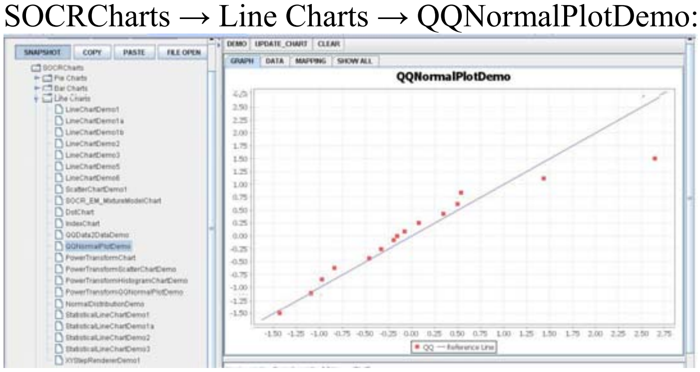


7. $P(155 \le X \le 186) = P(\frac{155-176}{30} \le Z \le \frac{186-176}{30}) = P(-0.7 \le Z \le 0.3333) = 0.3886$



Question 3. The June precipitation totals, in inches" for the city of Cleveland, OH are given below. Use these values to create a normal probability plot of the data. Do you consider these data to be Normally distributed?

Answer:
I used SOCR's QQNormalPlotDemo to produce the normal probability plot. You can find it under SOCRCharts → Line Charts → QQNormalPlotDemo:



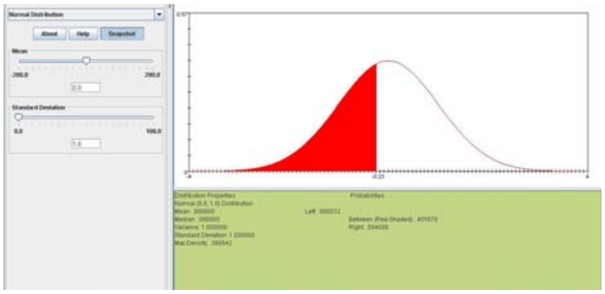
The data is approximately normal since all data points approximately lie on a straight line (though the right tail shows some deviation from the line).

Question 4. The litter size of a certain population of female mice follows approximately a normal distribution with mean 8.0 and standard deviation 2.1. Let Y be the size of a randomly chosen litter. Use the Normal Distribution to find approximate values for these probabilities:

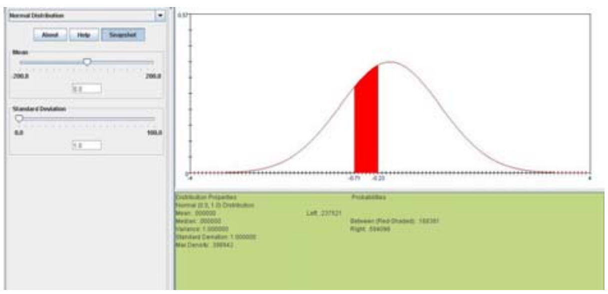
P(Y <= 7)
P(Y = 7)
P(6 <= Y <= 11)

Answer:
Note that the litter size is a discrete random variable. So here we only use normal distribution as an approximation, and we apply continuity correction in the computation. Without continuity correction, the results are acceptable, but are slightly different.

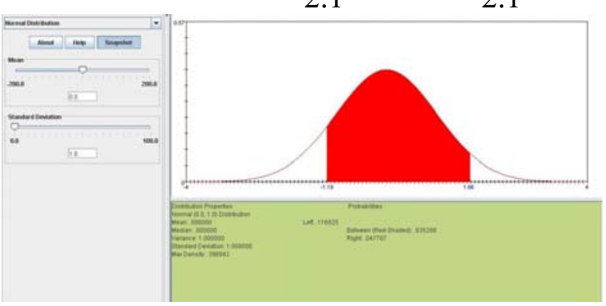
1. $P(Y \le 7) = P(Z \le \frac{7.5-8}{2.1}) = P(Z \le -0.2381) = 0.4059$



2. $$P(Y = 7) = P(Y \le 7) - P(Y \le 6) = P(Z \le \frac{7.5-8}{2.1}) - P(Z \le \frac{6.5-8}{2.1}) = P(Z \le -0.2381)$$
$$- P(Z \le -0.7143) = P(-0.7143 < Z \le -0.2381) = 0.1684$$



3. $P(6 \le Y \le 11) = P(\frac{5.5-8}{2.1} \le Z \le \frac{11.5-8}{2.1}) = P(-1.1905 \le Z \le 1.6667) = 0.8353$



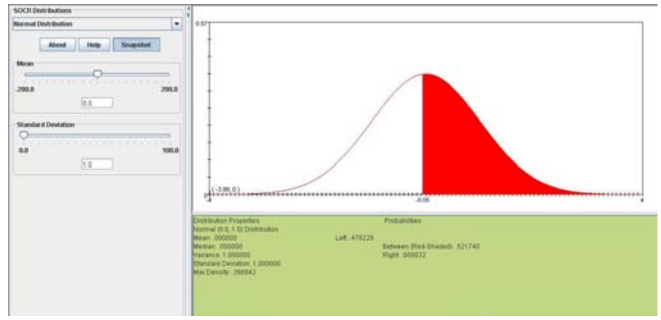
Question 5. A survey of mitochondrial DNA variation in smelts in a lake revealed that 2 haplotypes (genotypes) were present in the population. 30% of the fish were of haplotype A, and the remaining 70% were haplotype B. If we sample 400 fish from the lake, what is the probability that:

at least 120 are haplotype A?
at least 310 are haplotype B?
Between 125 and 145 are haplotype A?
170 or more are haplotype A?
Simulate these experiments using the SOCR Binomial Coin Experiment. Compare your exact calculations with the results of your simulations.

Answer:

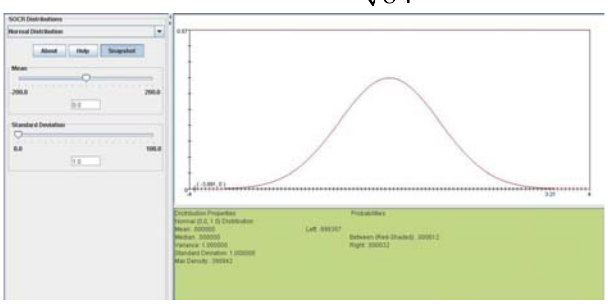
Let Y_A be the number of fish that are haplotype A, and Y_B be the number of fish that are haplotype B. Then $Y_A \sim$ Binom(400, 0.3), and $Y_B \sim$ Binom(400, 0.7).

This is a binomial problem. Since the sample size 400 is large enough, we can use normal approximation with continuity correction to compute these probabilities. Specifically, we use $N(400*0.3,\ \sqrt{400*0.3*0.7}) = N(120,\ \sqrt{84})$ to approximate the distribution of Y_A, and $N(400*0.7,\ \sqrt{400*0.3*0.7}) = N(280,\ \sqrt{84})$ to approximate the distribution of Y_B. Note that without continuity correction, the results are acceptable, but are slightly different.

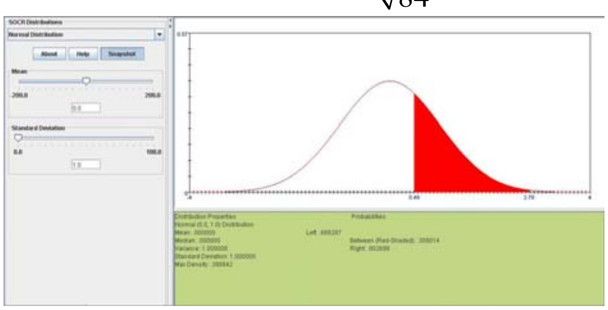


1. $P(Y_A \geq 120) = P(Z \geq \frac{119.5-120}{\sqrt{84}}) = P(Z \geq -0.0546) = 0.5218$

2. $P(Y_B \geq 310) = P(Z \geq \frac{309.5-280}{\sqrt{84}}) = P(Z \geq 3.2187) = 0.0006$



3. $P(125 \leq Y_A \leq 145) = P(\frac{124.5-120}{\sqrt{84}} \leq Z \leq \frac{145.5-120}{\sqrt{84}}) = P(0.4910 \leq Z \leq 2.7823) = 0.3090$



4. $P(Y_A \geq 170) = P(Z \geq \frac{169.5-120}{\sqrt{84}}) = P(Z \geq 5.4009) = 0$.

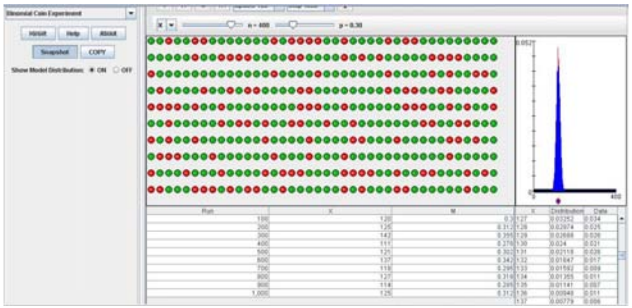
5. Simulation results: I ran 1000 simulations using SOCR's applet. Each of you should have different results based on your own simulations.

When I used n = 400, p = 0.3, I obtained the following result:

The empirical value of $P(Y_A \geq 120)$ is 0.505.

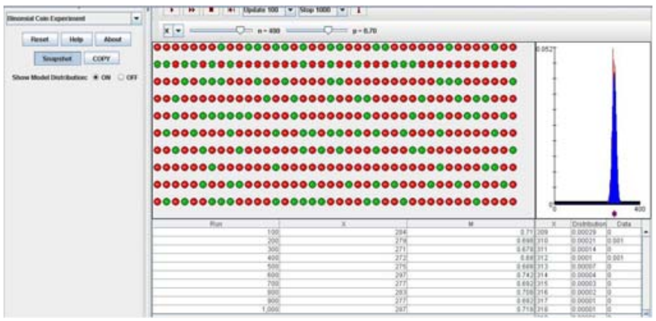
The empirical value of $P(125 \leq Y_A \leq 145)$ is 0.302.

The empirical value of $P(Y_A \geq 170)$ is 0.



When I used n = 400, p = 0.7, I obtained the following result:

The empirical value of $P(Y_B \geq 310)$ is 0.002.

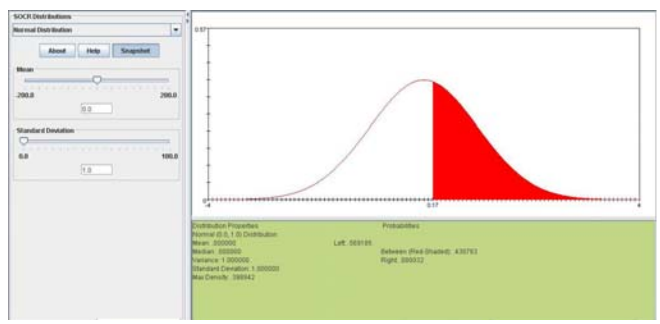


Question 6. Resting heart rate was measured for a group of subjects; the subjects then drank 6 ounces of coffee. Ten minutes later their heart rates were measured again. The change in heart rate followed a normal distribution, with mean increase of 7.1 beats per minute and a standard deviation of 10.9. Let Y denote the change in heart rate for a randomly selected person. Find:

P(Y > 9).

P(Y > 19).

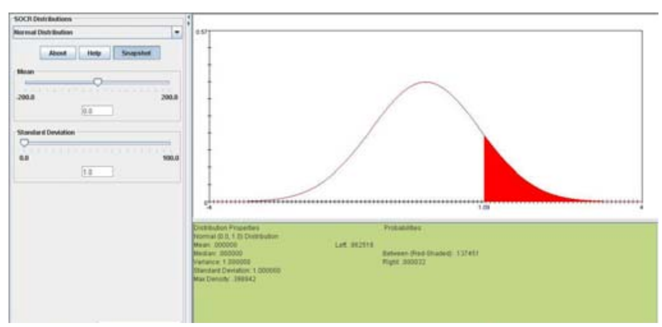
P(6 < Y < 14).



Answer:

1. $P(Y > 9) = P(Z > \frac{9 - 7.1}{10.9}) = P(Z > 0.1743) = 0.4308$

2. $P(Y > 19) = P(Z > \frac{19 - 7.1}{10.9}) = P(Z > 1.0917) = 0.1375$



3. $P(6 < Y < 14) = P(\frac{6-7.1}{10.9} < Z < \frac{14-7.1}{10.9}) = P(-0.1009 < Z < 0.6330) = 0.2768$