

1.

$$1)E(X) = 1/5(E(X_1)+ E(X_2)+ E(X_3)+ E(X_4)+ E(X_5)) = 3$$

$$SD(X) =$$

$$(1/25*SD(X_1)^2+1/25*SD(X_2)^2+1/25*SD(X_3)^2+1/25*SD(X_4)^2+1/25*SD(X_5)^2)^{(1/2)} =$$

$$1.3416$$

2)It is not possible to find the actual distribution and the $n = 5$ is too small to follow the normal distribution. It would not be a good estimate.

2.

$$P = 3.971503e-07$$

Consider the experiment in a total way. Y is the average of the purine amount. Which means the total amount of purine is greater or equal to 1500 in 2000 total experiment times. So it is $p = 0.7$ and size = 2000 while the set_x is from 1500 to 2000

3.

$$P=0.9623641 \ 0.9623659$$

Every time we generate 50 samples and we calculate the mean value of X and Y. if $\text{mean } X + 0.5 < \text{the mean of } Y$. Then we give the success count add 1. Each time we will run 10000 times to calculate the possibility. Then we repeat this for 20 time and calculate the 95% CI value.

4.

Mean value of 95% CI: 17.99271-18.22723

Var of 95% CI: 13579.09 58736.27

First generate x_1, x_2, x_3 using R built in function to generate 1000 rows sample, each row with 100000 sample. Then use them to calculate y. Then just calculate the mean and SD value with 95% CI. It looks like the var value is really big considering the sample size. It is mostly contribute by x_3^2 . Although t distribution has a mean of 0 but square of it is not.