# Stat13 Homework 6

http://www.stat.ucla.edu/~dinov/courses_students.html

**Suggested Solutions**

**HW_6_1** Let X = price of 100 g. X is approximately Normal.
Plan 1: T1 = 5 X
Plan 1: T2 = Y1 + Y2 + Y3 + Y4, where each Y is 1.25 X

Since X is Normal(mean = 1120, SD = 100), we have: Var(X) = 10000
E(Y) = 1.25 * 1120 = 1400, Var(Y) = 1.25 * 1.25 * Var(X) = 15625, SD(Y) = 125.
Y is Normal(mean = 1400, 125)

(a) E(T1) = 5 * 1120 = 5600.
Var(T1) = 25 * Var(X), so SD(T1) = sqrt(25 * 100 * 100) = 500.
T1 is Normal(mean = 5600, SD = 500).

(b) E(T2) = 4 * (1.25 * 1120) = 5600.
Var(T2) = Var(Y1 + Y2 + Y3 + Y4) = 4 *Var(Y1) = 4 * 15625, SD(T2) = 250.
Here we assume each Y's are independent.
T2 is Normal(mean = 5600, SD = 250).

(c) The variance of plan 1 is larger.

(d) Plan 1:
P(T1 > 5100) =
P( (T1 – 5600)/500 > (5100 – 5600)/500 ) = P(Z > -1) = 0.84

(e) Plan 2:
P(T2 > 5100) =
P( (T2 – 5600)/250 > (5100 – 5600)/250 ) = P(Z > -2) = .975

(f) Plan 1:
P(T1 < 4500) =
P( (T1 – 5600)/500 < (4500 – 5600)/500 ) = P(Z < -2.2) = 0.0139

(g) Either one of the following explanations is acceptable:
*Explanation 1*: Probability in (d) > Probability in (e), so it is more likely to exceed $5100 using plan 1. Therefore, plan 2 is safer.
*Explanation 2:* Smaller variance yields better prediction. Plan 1 has smaller variance so it is safer.

**HW_6_2**
Each die: Let X = possible outcomes, so X could take on 1, 2, 3, 4, 5, 6, 7, and 8.
P(X = 1) = P(X = 2) = …. = 1/8.

$E(X) = (1 + 2 + 3 + \ldots + 8) / 8 = 36/8 = 4.5$
$Var(X) = ( (1 - 4.5)^2 + (2 - 4.5)^2 + \ldots + (8 - 4.5)^2 ) / 8 = 6.$

Five dice is rolled twice --

Let $X_{i,j}$ = the outcome of the ith die, at the jth time, i = 1, 2, 3, 4, 5, and j = 1, 2. (This is just one way to assign subscripts.)

Then $Y = X_{1,1} + X_{1,2} + \ldots + X_{5,1} + X_{5,2}$ (there are 10 terms here.)

Note each $X_{i,j}$ is independent from another, so variance of the sum is the sum of the variance.

$m_Y = E(Y) = 2 * 5 * 4.5 = 45.$
$Var(Y) = Var(X_{1,1} + X_{1,2} + \ldots + X_{5,1} + X_{5,2}) = 2 * 5 * 6 = 60.$

So SD(Y) = sqrt(60) = 7.745967
Now we carry out the experiment 9 times:

Let $\bar{Y}$ = sample mean of five dice being rolled twice. Sample size = 9.
The by CLT, $\bar{Y}$ is Normal. To estimate the mean and SD of $\bar{Y}$:
Mean($\bar{Y}$) = 45
SD($\bar{Y}$) = sqrt( Var(Y) / 9) = 2.581989

**HW_6_3**

Let X = number of correctly remembered words of a mnemonics group subject. (like "treatment")
Let Y = number of correctly remembered words of a normal group subject. (like "control")

These two groups of samples are independent.

(a) For the sampled data:
mean(X) = 14.1, SD(X) = 2.468752
mean(Y) = 9.631579, SD(Y) = 3.33684

(b) Let $\bar{X}$ = sample mean of the normal group.
Let $\bar{Y}$ = sample mean of the normal group
An estimate of the "difference in the mean" is D = $\bar{X} - \bar{Y}$ = 4.468421

Sample variance of $\bar{X}$, $S_x^2 = 11.13450$
Sample variance of $\bar{Y}$, $S_y^2 = 6.094737$
$SD(D) = sqrt(S_x^2/19 + S_y^2/20) = 0.9438026$

And D follows a t-distribution of DF = min(20 –1, 19 – 1) = 18.
$t_{0.975} = 2.101$
95% CI = 4.468 ± 2.101 * 0.944 = (2.485492, 6.45135)

Plain English sentence: The 95%CI does not cover zero, and this suggests that the difference, D, is significantly different from zero.

(c) Approximate twice as much as old sample size.
We have: n1 = 20 and n2 = 19 are about the same. So, use n1 ≈ n2.
To find the new CI, plug in:

New n1 = 4 n1,
New n2 = 4 n2 ≈ 4 n1

Sample variance $S_x^2$ and $S_y^2$ are approximately the same as they were before plugging using 4 n1.

Then, the new CI = 0.5 old CI.

Note that $t_{0.975}$ does not change much (still about 2).
So the new sample should be four times of the old one.

(d) Since CI is 95% CI, we expect the CI to cover the true parameter value with 95% probability.