04b Sample Examination Problems Chapter 8 SOLUTIONS

1. The table below shows the annual salaries in dollars of randomly selected faculty in public educational institutions and private educational institutions.

Public	52127	57380	34122	8334	35730	22411	40196
Private	40807	26448	48970	52411	20223	39421	40102
Public	28528	10562	33666				
Private	46461	32557					

   (a) Find a 90% confidence interval for the difference between population mean annual salaries in the public and private institutions.

Although it looks like paired samples, it is not Since the sample sizes are different.

We need the 90% of $\mu_Y - \mu_X$

Public X : $\bar{X} = \dfrac{\sum_{i=1}^{10} X_i}{n} = \dfrac{323056}{10} = 32305.6$

$$S_X^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2 = 251149275$$

$$\Rightarrow S_X = 15847.7$$

Private Y : $\bar{Y} = \dfrac{\sum_{i=1}^{9} Y_i}{n} = \dfrac{347400}{9} = 38600$

$$S_Y^2 = \frac{1}{n-1}\sum_{i=1}^{n}(Y_i - \bar{Y})^2 = 111222489$$

$$\Rightarrow S_Y = 10546.2$$

Assumption : The population variances are unknown(not given)

The population variances are equal.

We use the pooled variance estimator:

$$S^2 = \frac{(n_x - 1)S_X^2 + (n_y - 1)S_Y^2}{n_x + n_y - 2} = \frac{9(251149275) + 8(111222489)}{17} = 185301375.7$$

$\Rightarrow S = \sqrt{S^2} = 13612.5$

The 90% C.I. , $\alpha = 0.1 \Rightarrow \alpha/2 = 0.05$ **, I will use** $\overline{Y} - \overline{X}$ **instead of** $\overline{X} - \overline{Y}$ **since** $\overline{Y}$ **= 38600 >** $\overline{X}$ **= 323056.5**

$$\overline{Y} - \overline{X} \pm t_{\alpha/2, n_X + n_Y - 2} S \sqrt{\frac{1}{n_x} + \frac{1}{n_y}} \quad , \quad t_{0.05,17} = 1.740$$

= 38600 – 32305.6 ± (1.740)(13612.5)$\sqrt{\frac{1}{10} + \frac{1}{9}}$

= 62944 ± 10882.9 = (-4588.5 , 17177.3)

That is we are 90% confident that $\mu_Y - \mu_X \in$ **(-4588.5 , 17177.3)**

<u>Remark</u>: Note that when 0 ∈ C.I. then we say at the 90% confidence level , the population means are potentially equal.

(b) Test the null hypothesis that mean salary for the private institutions is 1000 dollars more than in the public institutions against the alternative that the mean for the private institutions is more than 1000 dollars greater.

H$_0$ **:** $\mu_Y - \mu_X$ **= 1000** **again assuming** $\sigma_x^2 = \sigma_y^2 = \sigma^2$

H$_1$ **:** $\mu_Y - \mu_X$ **> 1000** **One tailed test (upper)**

The test statistics : $\dfrac{\overline{Y} - \overline{X} - (\mu_Y - \mu_X)}{S\sqrt{\frac{1}{n_x} + \frac{1}{n_y}}} \sim t_{n_X + n_Y - 2}$

Since we were not given α **.we'll use** α **= 0.05**

The critical value from table 10 : $t_{17,5} = 1.74$

$$\frac{\overline{Y} - \overline{X} - (\mu_Y - \mu_X)}{S\sqrt{\frac{1}{n_x} + \frac{1}{n_y}}} = \frac{38600 \;\; 32305.6 - (1000)}{13612.5\sqrt{\frac{1}{10} + \frac{1}{9}}} = 0.8465 \prec 1.74$$

It falls outside the rejection region and therefore we do Not Reject H$_0$ **i.e. there is no significant evidence that the difference between private and public sectors salaries exceeds 1000.**

(c) State carefully the assumptions you have made in arriving at the test and confidence interval.

The assumptions :

-Normality : Normally distributed salaries in the private and the public sectors

$$\overline{X} \sim N(\mu_x, \frac{\sigma_x^2}{n_x}) \ , \ \overline{Y} \sim N(\mu_y, \frac{\sigma_y^2}{n_y})$$

-Independence :so we were able to use: 0

$$Var(\overline{Y} - \overline{X}) = Var(\overline{Y}) + Var(\overline{X}) - 2Cov(\overline{X}, \overline{Y}) = \frac{\sigma_x^2}{n_x} + \frac{\sigma_y^2}{n_y}$$

- Main assumption : $\sigma_x^2 = \sigma_y^2 = \sigma^2$ **and unknown.**

---

2. Primary school children with reading problems were randomly divided into a control group and a group that received special reading teaching. The results of a subsequent reading test for all the children are given below:

Control	42	43	55	26	62	37	33	41	19	54	20	
	85	46	10	17	60	53	42	37	42	55	28	48
Special	24	43	58	71	43	49	61	44	67	49		
Teaching	53	56	59	52	62	54	57	33	46	43	57	

(a) Find a 99% confidence interval for the difference in score between the controls and the specially-taught group.

We need 99% C.I. for $\mu_Y - \mu_X$

Control X : $$\bar{X} = \frac{\sum_{i=1}^{23} X_i}{n} = 41.52$$

$$S_X^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2 = 294.08 \Rightarrow S_X = 17.15$$

Special Y : $$\bar{Y} = \frac{\sum_{i=1}^{21} Y_i}{n} = 51.48$$

$$S_Y^2 = \frac{1}{n-1}\sum_{i=1}^{n}(Y_i - \overline{Y})^2 = 121.16 \Rightarrow S_Y = 11.01$$

$n_x = 23$, $n_y = 21$

We use the pooled variance estimator:

$$S^2 = \frac{(n_x - 1)S_X^2 + (n_y - 1)S_Y^2}{n_x + n_y - 2} = \frac{22(294.08) + 20(121.16)}{42} = 211.737$$

$$\Rightarrow S = \sqrt{S^2} = 14.55$$

The 99% C.I. , $\alpha = 0.01 \Rightarrow \alpha/2 = 0.005$ **, I will use** $\overline{Y} - \overline{X}$ **instead of** $\overline{X} - \overline{Y}$ **since** $\overline{Y}$ **= 51.48 >** $\overline{X}$ **= 41.52**

$\overline{Y} - \overline{X} \pm t_{\alpha/2, n_X + n_Y - 2} S\sqrt{\frac{1}{n_x} + \frac{1}{n_y}}$ **,** $t_{0.005,42} \approx 2.704$ **(using 40 df)**

= 51.48 – 41.52 $\pm$ **(2.704)(14.55)** $\sqrt{\frac{1}{23} + \frac{1}{21}}$

= 9.96 $\pm$ **11.96 = (-1.915 , 21.835)**

That is we are 99% confident that $\mu_Y - \mu_X \in$ **(-1.915 , 21.835)**

<u>Remark</u>: Note that when 0 $\in$ **C.I. then we say at the 99% confidence level , the population means are potentially equal. If you find the 90% C.I , the interval becomes (2.56 , 17.36) which excludes 0 and we say: there is a significant difference between the population means.**

(b) Test at the 10% level the null hypothesis that there is no difference between the two groups.

H_0 : $\mu_Y = \mu_X$ **again assuming** $\sigma_x^2 = \sigma_y^2 = \sigma^2$

H_1 : $\mu_Y \neq \mu_X$ **Two tailed test**

The test statistics : $\dfrac{\overline{Y} - \overline{X} - (\mu_Y - \mu_X)}{S\sqrt{\frac{1}{n_x} + \frac{1}{n_y}}} \sim t_{n_X + n_Y - 2}$

α **=0.1 ,** $\alpha/2$ **= 0.05**

The critical value from table 10 : $t_{42,5} = 1.648$

$$\frac{\overline{Y} - \overline{X} - (\mu_Y - \mu_X)}{S\sqrt{\frac{1}{n_x} + \frac{1}{n_y}}} = \frac{9.96 - (0)}{14.55\sqrt{\frac{1}{23} + \frac{1}{21}}} = 2.268 \succ 1.684$$

It falls within the rejection region and therefore we Reject H_0 i.e. there is significant difference between special and control scores.