```
[1] -0.75 -1.50  3.75 -6.75 -1.50 -2.25 -0.75  4.50  3.75 -1.50 -2.25  3.00  2.25 -3.00
3.00
> trtN <- c(10,8,16,15)
> trtS <- c(9,8)
> mean(trtN)
[1] 12.25
> mean(trtS)
[1] 8.5
```

(25 marks, 5 marks each)

a) State the null and alternative hypotheses of this experiment.

5

μ_N - mean survival for treatment N

μ_S - mean survival for treatment S

$H_0: \mu_N = \mu_S$ vs. $H_A: \mu_N \neq \mu_S$

b) What is the probability that a patient will receive the new treatment? Does the observed data make you suspicious that the coin used was not fair? Explain.

Since the coin is fair $P(Trt = N) = P(Trt = S) = 1/2$

The coin is fair even though the observed treatment assignment was NNN SS N. This has same probability as the assignment NNN SSS using a fair coin.

c) List all the values of the randomization distribution of the mean difference.

-0.75	-1.50	3.75	2.25
-1.50	-2.25	-1.50	-3.00
3.75	-0.75	-2.25	
-6.75	4.50	3.00	

d) **What is the empirical cumulative distribution function of the randomization distribution?**

$\hat{F}(x) = \#\{\text{Values} < x\} / 20$

Where Values are from part (c).

e) Is there evidence that the difference in means is due to random chance? If you ~~didn't~~ did find evidence state two possible reasons?

$$\text{P-value} = 2\left(1 - \hat{F}\left(12.25 - 8.5\right)\right)$$

$$= 2\left(1 - \hat{F}\left(3.75\right)\right)$$

$$= 2\,\#\{|\text{Values}| \geq 3.75\}/20$$

$$= 2\left(4/20\right) = 8/20 = 0.4$$

There is evidence that the difference is due to chance.

Possible reasons: 1 Low power

2 No real difference between two treatments

a) Describe the type of experimental design used in this experiment. What are the experimental units? What are the treatments?

- paired design
- experimentall units are people
- treatments are nalbuphione and morphine.

b) Is there statistical evidence at the 10% significance level that the treatments are different? If there is evidence of a difference then which treatment resulted in a larger mean pupil diameter?

paired t-test has p-value = 0.071 < 0.10

∴ there is evidence at 10% sig. level that treatments are different.

Nalbuphione resulted in larger pupil diameter. (treatment A)

c) For your answer in part b) what statistical assumptions does your answer rely on?

1. Differences within each Subject are normally distributed

d) Describe how you would conduct a randomization test to evaluate the evidence against the null hypothesis. Would you expect that the results would be the same as the t-test?

1. There are $2^{12} = 4096$ possible outcomes under H_0.
2. Compare average difference 0.747 with the 4095 average differences that could have occured as a result of different coin tosses.

e) Suppose that another scientist wanted to design another similar study that will have at least 80% power to detect a difference of 0.5mm at the 5% significance level. How many patients should be assigned to each treatment group? Would you expect the number of patients required for 80% power to increase or decrease if the scientist changed her mind and instead wanted to detect a difference of 0.4mm? Explain.

- Based on power.t.test for paired design.

  $n \approx 55$

- Would expect number of patients to increase

  $\because$ want to detect smaller effect. i.e.,

  $\delta < 0.5$ mm $\Rightarrow$ $n > 55$ holding other

  parameters constant.

3. Consider the ANOVA model $y_{it} = \mu + \tau_t + \epsilon_{it}$, where $i = 1,..,n$, $t = 1,..,v$, $\varepsilon_{it} \overset{iid}{\sim} N(0,\sigma^2)$, and $\sum_{t=1}^{v} \tau_t = 0$. **(25 marks)**

a) Briefly explain what the parameters μ, τ_t, and ϵ_{it} represent. (3 marks)

$\mu =$ grand mean

$\tau_t =$ mean of treatment $t = 1, \ldots, v$

$\varepsilon_{it} =$ within treatment variability.

b) What is the ANOVA model if all the treatment means are equal? (2 marks)

$\tau_1 = \tau_2 = \cdots = \tau_t = \tau.$

$\therefore \quad y_{it} = \mu + \tau + \varepsilon_{it}, \quad \varepsilon_{it} \sim N(0, \sigma^2)$

c) Find the least squares estimates of τ_t, $t = 1, \ldots, v$. (10 marks)

$$\sum_{i=1}^{n} \sum_{t=1}^{v} \varepsilon_{it}^2 = \sum_i \sum_t (y_{it} - \mu - \tau_t)^2 = L(\mu, \tau_1, \ldots, \tau_v)$$

$$\left.\frac{\partial L}{\partial \mu}\right|_{(\hat\mu, \hat\tau_i)} = 0 \quad , \quad \left.\frac{\partial L}{\partial \tau_i}\right|_{(\hat\mu, \hat\tau_i)} = 0$$

$$\sum_i \sum_t 2 (y_{it} - \mu - \tau_t)(-1) = 0 \qquad \sum_i \left[(y_{i1} - \mu - \tau_1)^2 + \cdots + (y_{iv} - \mu - \tau_v)^2 \right]$$

$$\Rightarrow \sum_i y_{i\cdot} - v\mu = 0 \qquad (1)\ \frac{\partial L}{\partial \tau_1} = \sum_i (y_{i1} - \mu - \tau_1) = 0$$

$$\vdots$$

$$\Rightarrow \hat\mu = y_{\cdot\cdot}/v \cdot n \qquad (2)\ \frac{\partial L}{\partial \tau_v} = \sum_i (y_{iv} - \mu - \tau_v) = 0$$

$$= \bar{y}_{\cdot\cdot}$$

$$\hat\tau_v$$

$$(1)\ \sum_i y_{i1} - n\hat\mu - n\hat\tau_1 = 0$$

$$\Rightarrow \hat\tau_1 = \frac{\sum y_{i1} - n\hat\mu}{n} = \bar{y}_{\cdot 1} - \hat\mu$$

$$(2)\ \sum_i y_{iv} - n\hat\mu - n\hat\tau_v = 0$$

$$\Rightarrow \hat\tau_v = \bar{y}_{\cdot 1} - \hat\mu$$

d) Show that the mean square for error $MSE = \frac{\sum_{i=1}\sum_{t=1}(y_{it} - \bar{y}_{i.})^2}{N-v}$, $N = nv$ is an unbiased estimator of σ^2 when all the treatment means are equal. (10 marks)

$$y_{it} = \mu + \tau_i + \varepsilon_{it}, \quad \varepsilon_{it} \sim N(0, \sigma^2)$$

$$\Rightarrow y_{it} \sim N(\mu + \tau_i, \sigma^2), \quad \bar{y}_{i.} \sim N(\mu + \tau_i, \sigma^2/n)$$

$$\sigma^2 = E(y_{it}^2) - (\mu + \tau_i)^2 \qquad \frac{\sigma^2}{n} = E(\bar{y}_{i.}^2) - (\mu + \tau_i)^2$$

$$SSE = \sum_i \sum_t \left(y_{it}^2 - 2 y_{it} \bar{y}_{i.} + (\bar{y}_{i.})^2\right)$$

$$= \sum_i \left(\sum_t y_{it}^2 - 2 \sum_t y_{it} \bar{y}_{i.} + \sum_t \bar{y}_{i.}^2\right)$$

$$= \sum_i \left(\sum_t y_{it}^2\right) - 2\bar{y}_{i.} n \bar{y}_{i.} + n\bar{y}_{i.}^2$$

$$= \sum_i \left(\sum_t y_{it}^2 - n\bar{y}_{i.}^2\right)$$

$$\therefore E(SSE) = \sum_i \left(\sum_t E(y_{it}^2) - n E(\bar{y}_{i.}^2)\right)$$

$$= \sum_i \left(n(\sigma^2 + (\mu + \tau_i)^2) - n\left(\frac{\sigma^2}{n} + (\mu + \tau_i)^2\right)\right)$$

$$= \sum_i (n-1)\sigma^2 = v(n-1)\sigma^2$$

$$= (vn - v)\sigma^2$$

$$= (N - v)\sigma^2$$

$$\therefore E(MSE) = E\left(\frac{SSE}{n - v}\right) = \sigma^2 \quad \boxed{\checkmark}$$

a) State the null and alternative hypo[illegible]s of the F test. Fill in the six values (??) of the ANOVA table below. Is t[illegible] evidence against the null hypothesis (the appropriate F critical value – [illegible] 5% point is 2.95)? (10 marks)

use 5% level.

```
> anova(ped.aov)
Analysis of Variance Table

Response: TIME
                    Df   Sum Sq    Mea[illegible]    F value Pr(>F)
as.factor(PUSHES) ?? 0.008047       ?[illegible]      ??       ??
Residuals          ?? 0.305953       ?[illegible]
```

	Df	SS	MS	F
Treatment	3	0.008047	0.0027	0.247
Residuals	28	0.305953	0.0109	
Total	31			

$P(F_{3,28,0.05} > 0.247)$

No evidence against Ho at 5% level.

b) State the assumptions of the F test. Are the assumptions satisfied? (5 marks)

1. $\varepsilon_{it} \overset{iid}{\sim} N(0, \sigma^2)$

2. additive model appropriate

Based on residual plot and normal QQ plot assumptions are satisfied.

c) If all possible pairs of mean times are compared using a type I error rate of 5% then are all pairwise confidence intervals 95% confidence intervals? Explain. **(5 marks)**

NO. Since P(at least one H_0 false | H_0 true) = FWER

> 0.05

All pairwise confidence intervals use an experiment-wise error rate of 5% which is different from the FWER.

d) Does this experiment support the following statement: "pushing the button has a small effect on the time a person has to wait for the "walk" sign to appear". State if you agree or disagree and justify your answer. **(5 marks)**

Disagree. There is no evidence that the difference in wait times between number of button pushes is not due to chance. The P-value of the F-test >0.05 and the smallest mean wait time is 38.17s and the largest is 38.12s. A difference of 38.17s - 38.12s = 0.05s does not seem to be practically significant.