STATISTICS FOR MANAGEMENT (IDS 570)

HOMEWORK 9 SOLUTION

Problem 1

Person	1	2	3	4	5	6	7	8	$\bar{x}$	stdv
Eye affected by glaucoma (x_1)	488	478	480	426	440	410	458	460	455	27.7
Eye not affected by glaucoma (x_2)	484	478	492	444	436	398	464	476	459	31.31
Difference ($x_1 - x_2$)	4	0	-12	-18	4	12	-6	-16	-4	10.744

(a) Measurements are taken of pairs of eyes, one with glaucoma and the other without, matched by the property of belonging to the same person. Since the population standard deviation is unkown, we use the paired T-test to anlyze this hypothesis test.

(b) Let $\bar{x}_1 = 455$ and $\bar{x}_2 = 459$ denote the mean corneal thickness of effected eye and not effected eye, repectively and s_1 and s_2 denote the corresponding standard deviations. The t-score can be computed using the following formula:

$$t = \frac{\bar{d} - 0}{s_{\bar{x}}/\sqrt{n}} = \frac{-4 - 0}{10.744/\sqrt{8}} = -1.053$$

where, $\bar{d} = \bar{x}_1 - \bar{x}_2$, and $s_{\bar{x}}$ is the standard deviaiton of the difference between x_1 and x_2.

To compute t-test, we can use the following R code:
> affected = c(488, 478, 480, 426, 440, 410, 458, 460)
> notaffected = c(484, 478, 492, 444, 436, 398, 464, 476)
> t.test(affected, notaffected, alternative="less", paired=TRUE)

(c) Using the t-distribution table, we get p-value is greater than 0.1 for degree of freedom 8-1 = 7. In addition, using the software R, we get p-value = 0.1637.

```
 Paired t-test

data:  x and y
t = -1.053, df = 7, p-value = 0.1637
alternative hypothesis: true difference in means is less than 0
95 percent confidence interval:
     -Inf 3.196549
sample estimates:
mean of the differences
                     -4
```

(d) Since the pvalue = 0.1637 > α = 0.1, we do not have enough evidence to support the claim that the average corneal thicknesses are affected by glaucome. Therefore, we conclude that the corneal thickness of the eye affected by glaucoma is the same as the corneal thickness of the other eye not affected by glaucoma.

(e) A one-sided 90% CI for $\mu_1 - \mu_2$ is given by

$$\left[-\infty, \bar{d} + t^* s_{\bar{x}}/\sqrt{n}\right] = \left[-\infty, -4 + 1.415 \times 10.744/\sqrt{8}\right]$$

That is, $[-\infty, 1.374]$.

To answer this question, we can also use the following R code:

> t.test(x, y, alternative="less", paired=TRUE , conf.level=0.90)

```
 Paired t-test

data:  x and y
t = -1.053, df = 7, p-value = 0.1637
alternative hypothesis: true difference in means is less than 0
90 percent confidence interval:
     -Inf 1.374583
sample estimates:
mean of the differences
                     -4
```

The null-hypothesis states that the difference between the means of the two samples is 0. Since 0 lies within the confidence interval, we do not reject the null hypothesis.

Problem 2

(a) To get a summey of vector x, we use the code "summary(x)" in R. Summary of data for children is

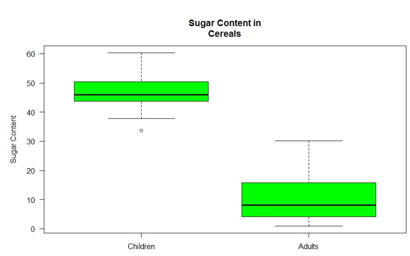
```
 Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
33.60   43.65   45.90   46.80   50.35   60.30
```

Summary of data for adults is

```
 Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
 1.00    4.10    8.10   10.17   15.80   30.20
```

(b) We use the following R code:

> boxplot(children, adults, names = c("Children", "Adults"), main = "sugar Contect in Cereals", xlab = "Cereals", ylab = "Suger Content", col = "green", las = 1)



(c) A 95% confidence interval for the different in mean sugar content in cereal for children and adult can be calculated using the following formula:

$$(\bar{x}_1 - \bar{x}_2) \pm t^* \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

where $\bar{x}_1$ and $\bar{x}_2$ denote the mean sugar contect in cereals for children and adults respectively and s_1 and s_2 are the corresponding standard deviations. Therefore the confidence interval is [32.5575, 40.7045]

Problem 3

(a) Let μ_1 and μ_2 denote the mean gas millage provided by premium gas and regular gas respectively. The test hypothesis is

$$H_0 : \mu_1 = \mu_2, \quad H_a : \mu_1 > \mu_2.$$

Since the same sample of the cars are used to run the test with regular gas and premium gas we use the paired T test.

(b) Summary of regular gas data is

```
 Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
16.00   21.25   22.50   23.10   26.50   28.00
```

Summary of premium gas data is

```
 Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
 19.0    24.0    25.0    25.1    26.0    32.0
```

(c) Use the following R code:

> t.test(premium, regular, alternative = “greater”, paired = TRUE) Using the R output, we

```
 Paired t-test

data:  pre and reg
t = 4.4721, df = 9, p-value = 0.0007749
alternative hypothesis: true difference in means is greater than 0
95 percent confidence interval:
 1.180207      Inf
sample estimates:
mean of the differences
                      2
```

have p-value = 0.0007749.

Problem 4

(a) The 98% confidence interval is

$$\bar{x} \pm t^* \frac{\sigma}{\sqrt{n}} = 98.2846 \pm 2.4017 \frac{0.6824}{\sqrt{52}} = [98.05732, 98.51188]$$

(b) The hypothesis testing is

$$H_0 : \mu = 98.6 \quad H_a : \mu \neq 98.6$$

Assuming H_0 is true, the t-score is $t = \dfrac{\bar{x} - \mu}{\sigma/\sqrt{n}} = \dfrac{98.2846 - 98.6}{\frac{0.6824}{\sqrt{52}}} = -3.333$. Then the p-value for the two-sided test is $2 \times P(T \leq -3.333) = 2 \times 0.0008 = 0.0016$. Since the p-value is less than α, we reject the null-hypothesis. This means that the mean body temperature is not equal to 98.6 °Fahrenheit.

(d) To find the power of the test, we can use the following R code:
> power.t.test(n = 52, delta = (98.2 − 98.2846), sd = 0.6824, sig.level =0.02, power = NULL, type = "one.sample", alternative = "two.sided",strict = TRUE)

```
     One-sample t test power calculation

              n = 52
          delta = 0.0846
             sd = 0.6824
      sig.level = 0.02
          power = 0.07341512
    alternative = two.sided
```

Problem 5

To answer this, we first need to state the hypotheses. The parameter of interest is p, the proportion of people who would receive superior UVA and UVB protection from your product. If there is no difference between these two lotions, then we'd expect $p = 0.5$. (In other words, which product works better on someone is like flipping a fair coin.) This is our null hypothesis. As for Ha, we'll use the two-sided alternative even though your company hopes for $p > 0.5$.

So, to answer this claim, we have $n = 20$ subjects and $X = 13$ successes and want to test

$$H_0 : p = 0.5, \quad H_a : p \neq 0.5$$

The expected numbers of successes (your product provides better protection) and failures (your competitor's product provides better protection) are $20 \times 0.5 = 10$ and $20 \times 0.5 = 10$. Both are at least 10, so we can use the z test. The sample proportion is

$$\hat{p} = \frac{X}{n} = \frac{13}{20} = 0.65$$

The test statistic is

$$z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} = \frac{0.65 - 0.5}{\sqrt{\frac{0.5(0.5)}{20}}} = 1.34$$

From Table A we find $P(Z \geq 1.34) = 0.9099$, so the probability in the upper tail is $1 - 0.9099 = 0.0901$. The P-value is the area in both tails, so $P = 2 \times 0.0901 = 0.1802$. We conclude that the sunblock-testing data are compatible with the hypothesis of no difference between your product and your competitors ($\hat{p} = 0.65, z = 1.34, P = 0.18$). The data do not provide you with a basis to support your advertising claim.