Semester 1 | Solution to Tutorial Set 9 | 2013

1. Answer: (e) Use t-table with df=25-1=24 and upper area=(1-0.9)/2=0.05.

2. Answer: (a) The 90% CI for μ: $\bar{x} \mp t_{n-1,\alpha/2}\frac{s}{\sqrt{n}} = 42.7 \mp 1.711\frac{5}{\sqrt{25}}$

3. Answer: (d) Reason: $Z = \frac{\hat{p}-p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} \sim N(0,1)$ by the CLT on the sample mean which is also the sample proportion $\hat{p} = X/n$ when the sample size n is large and $p = p_0$ under H_0.
   Note that (c) $X \sim B(n, p_0)$ is also right. However the exact P-value based on binomial distribution is more complicated to calculate when n is large but it can be easily obtained from `R`.

4. Answer: (a) $Z = \dfrac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} = \dfrac{\frac{36}{200} - 0.24}{\sqrt{\frac{0.24(1-0.24)}{200}}} = -1.99$

5. Answer: (d) Since the P-value$=P(Z < -1.99) = 1 - 0.9767 = 0.0233 > 0.01$, the data are consistent with H_0 that the proportion of success is 24%.

6. The 99% CI for p: $\hat{p} \mp z_{1-\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = 0.18 \mp 2.575\sqrt{\frac{0.18(1-0.18)}{200}} = 0.18 \mp 0.07 = (0.11, 0.25)$ since $\hat{p} = \frac{36}{200} = 0.18$.

   Note that we do NOT check if $p_0 = 0.24 \in (0.11, 0.25)$ as an alternative procedure for testing $H_0 : p = 0.24$ in Q3-5 because the SE in the CI is calculated using $\hat{p} = 0.18$ instead of $p_0 = 0.24$ under H_0.

7. We have $n = 8$, $\bar{d} = 1.14$ and $s_d = 12.22$. The paired t-test on the difference between two means is:
   1. Hypotheses: $H_0 : \mu_d = 0$ against $H_1 : \mu_d \neq 0$
   2. Test statistic: $t_{\text{obs}} = \frac{\bar{d}-0}{s_d/\sqrt{n}} = \frac{1.14-0}{12.22/\sqrt{8}} = 0.264$
   3. P-value: $2P(t_{\text{obs}} > 0.264) > 2(0.25) = 0.5$.
   4. Conclusion: Since the P-value > 0.05, the data are consistent with H_0 that the dietary counseling made no difference in reducing sodium intake.

8. Use `pnorm((36/200-0.24)/sqrt(0.24*(1-0.24)/200))`. Answer is 0.02347236.
   Use `2*(1-pt(1.14/(12.22/sqrt(8)),7))`. Answer is 0.7994847.

9.
```
> diff=scan()
1: -1.74 -22.47 17.29 -6.84 4.99 9.27 9.51 -0.88
9:
Read 8 items
> t.test(diff)

        One Sample t-test

data:  diff
t = 0.2642, df = 7, p-value = 0.7993
alternative hypothesis: true mean is not equal to 0
95 percent confidence interval:
 -9.074864 11.357364
```

Note that the 95% CI for μ_d is (-9.074864, 11.357364). Since it includes $\mu_d = 0$, μ_d is not significantly different from zero. Besides the P-value=0.7993$>$ 0.05. Hence the conclusion is still "the data are consistent with H_0".

10.
```
> x=scan()
1: 7.85 12.03 21.84 13.94 16.68 41.78 14.97 12.07
9:
Read 8 items
> y=scan()
1: 9.59 34.50 4.55 20.78 11.69 32.51 5.46 12.95
9:
Read 8 items
> t.test(x, y, paired = T)
```

Result same as in Q10.

11.
```
> lead=scan(file=url("http://www.maths.usyd.edu.au/math1015/r/lead.txt"))
Read 124 items
> lead
  [1] 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
 [38] 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 1 1 1 1 1 1 1
 [75] 1 1 1 1 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
[112] 0 1 1 1 1 1 0 0 0 0 0 0 0
> n=length(lead)
> x=length(lead[lead==1])
```

```
> ph=x/n
> p0=0.15
> z0=(ph-p0)/sqrt(p0*(1-p0)/n)
> pvalue=2*(1-pnorm(z0))
> c(x,n,ph,z0,pvalue)
[1]  24.0000000 124.0000000   0.1935484   1.3580877   0.1744358
```

The count `x` is 24.
The sample proportion is $\frac{24}{124} = 0.1935$.
The test statistic is $Z_0 = \frac{\hat{p}-p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} = \frac{\frac{24}{124}-0.15}{\sqrt{\frac{0.15(1-0.15)}{124}}} = 1.3581$
The P-value is $2P(Z > 1.3581) = 0.1744$.
Since the P-value > 0.05, the data are consistent with $H_0 : p = 0.15$.

Additional Problems to Week 9 - Solutions

1. We have $\mu_0 = 1800$, $n = 50$, $\bar{x} = 1850$ and $s = 100$. The one sample t-test on the true mean μ is:

   1. Hypotheses: $H_0 : \mu = 1800$ against $H_1 : \mu > 1800$

   2. Test statistic: $T_{49} = \dfrac{\bar{x} - \mu_0}{s/\sqrt{n}} = \dfrac{1850 - 1800}{50/\sqrt{50}} = 7.071$

   3. P-value: $P(t_{49} > 7.071) < 0.001$.

   4. Conclusion: Since the P-value < 0.01, there is strong evidence in the data against H_0. The breaking strength of the steel wires produced by the new technique is higher than 1800N.

2. We have $n = 1000$, $\hat{p} = \frac{550}{1000} = 0.55$ and $p_0 = 0.5$. The one sample Z-test on the true proportion p is:

   1. Hypotheses: $H_0 : p = 0.5$ against $H_1 : p > 0.5$

   2. Test statistic: $Z = \dfrac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} = \dfrac{0.55 - 0.5}{\sqrt{\frac{0.5(1-0.5)}{1000}}} = 3.16$

   3. P-value: $P(Z > 3.16) = 1 - 0.9992 = 0.0008$

   4. Conclusion: Since the P-value < 0.05, there is strong evidence in the data against H_0. Candidate A will receive more than 50% of the vote.

3. We have $n = 10$, $\bar{d} = -4$ and $s_d = 5.7735$.

Twin	1	2	3	4	5	6	7	8	9	10
First	32	36	21	30	49	27	39	38	56	44
Second	44	43	28	39	51	25	32	42	64	44
Diff	-12	-7	-7	-9	-2	2	7	-4	-8	0

The paired sample t-test to compare two means is:

1. Hypotheses: $H_0 : \mu_d = 0$ against $H_1 : \mu_d \neq 0$

2. Test statistic: $T_9 = \dfrac{\bar{d} - 0}{s_d/\sqrt{n}} = \dfrac{-4 - 0}{5.7735/\sqrt{10}} = -2.19$

3. P-value: $2P(t_9 < -2.19) > 2(0.25) = 0.5$

4. Conclusion: Since the P-value > 0.05, the data is consistent with H_0 that the annual incomes for the first and second born twins are equal.