Answer Key (white paper)

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
C	B	A	C	A	C	B	C	B	B	D	B	B	A	D	D

1. The degrees of freedom associated with a chi-square test for two-way tables with *r* rows and *c* columns is
   A) $r - 1$
   B) $c - 1$
   C) $(r - 1) \times (c - 1)$
   D) $r \times c - 1$

2. In a two-tailed test for the population mean, if the null hypothesis is rejected when the alternative hypothesis is true,
   A) a Type II error is committed.
   B) a correct decision is made.
   C) a one-tailed test should be used instead of a two-tail test.
   D) a Type I error is committed.

Solution: *if we rejected H_0 when it is true, we would commit the type I error (this probability is equal to the level of significance α), but if we rejected the false null hypothesis, a correct decision has been made.*

Use the following to answer questions 3-4:

A manufacturer receives parts from two suppliers. A simple random sample of *400* parts from supplier A finds *25* defective. A simple random sample of *200* parts from supplier B finds *15* defective. Let p_1 and p_2 be the proportion of all parts from suppliers A and B, respectively, that are defective. Suppose that we are interested in testing the null hypothesis $H_0: p_1 - p_2 = 0$ using a two-tailed test at the *10%* significance level.

3. What is the estimate for $p_1 - p_2$, the difference in proportions?
   A) –0.0125
   B) 0.0067
   C) – 0.0067
   D) 0.0125

Solution: $\hat{p}_1 - \hat{p}_2 = \frac{25}{400} - \frac{15}{200} = -0.0125$

4. The critical values for a two-tailed test at the *10%* level of significance are $z_{crit} = -1.645$ and $z_{crit} = 1.645$. Which statement is true?
   A) The test statistic $z = -0.58$ is between the critical values, so H_0 should be rejected.
   B) The test statistic $z = -0.58$ is not between the critical values, so H_0 should be rejected.
   C) The test statistic $z = -0.58$ is between the critical values, so H_0 should not be rejected.
   D) The test statistic $z = -0.58$ is not between the critical values, so H_0 should not be rejected.

<u>Solution:</u> *the test statistic is between the critical values for the two-tailed test:* $z_{crit} = -1.645 < z = -0.58 < 1.645$. *Therefore,* H_0 *should not be rejected.*

---

5. When calculating a one-sample *t* confidence interval, which level of confidence will give the smallest interval?
   A) *90%*
   B) *95%*
   C) *96%*
   D) *99%*

6. A simple random sample of *100* postal employees is used to test if the average time postal employees have worked for the postal service has changed from the value of *7.5* years recorded 20 years ago. The sample mean was $\bar{x} = 7.1$ years with a standard deviation of $s = 2$ years. What is a *90%* confidence interval for μ?
   A) $7.1 - 2 < \mu < 7.1 + 2$
   B) $7.1 - 0.202 < \mu < 7.1 + 0.202$
   C) $7.1 - 0.332 < \mu < 7.1 + 0.332$
   D) $7.1 - 0.396 < \mu < 7.1 + 0.396$

<u>Solution:</u> *there is the confidence interval for the population mean μ with unknown standard deviation σ:*

$\bar{x} - t^* \times \frac{s}{\sqrt{n}} < \mu < \bar{x} + t^* \times \frac{s}{\sqrt{n}}$. *The critical value* $t^* = t(df = 99, c = 0.9) = 1.66$. *Therefore,*

$$7.1 - 1.66 \times \frac{2}{\sqrt{100}} < \mu < 7.1 + 1.66 \times \frac{2}{\sqrt{100}} \Leftrightarrow 7.1 - 0.332 < \mu < 7.1 + 0.332$$

7. Suppose two simple random samples of size *75* and *25* are drawn from a Normal population. What degrees of freedom could be used to perform a two-sample *t* test?
   A) *74*
   B) *24*
   C) *75*
   D) *25*

<u>Solution:</u> $df = \min(n_1, n_2) - 1 = \min(25, 75) - 1 = 25 - 1 = 24$.

8. A simple random sample of *160* blood donors is taken to estimate the proportion of donors with type A blood with a *95%* confidence interval. In the sample, there are *20* people with type A blood. What is the margin of error for this confidence interval?
   A) *0.0738*
   B) *0.0261*
   C) *0.0512*
   D) *1.96*

<u>Solution:</u> It is the margin of error of the single population proportion: $m = z^* \times \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$. *The critical value can be found at the t-distribution table: z*=1.96, the sample proportion:* $\hat{p} = \frac{20}{160} = 0.125$. *Putting the numbers into the formula, we will get:* $m = z^* \times \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = 1.96 \times \sqrt{\frac{0.125(1-0.125)}{160}} = 0.0512$

9. For a *99%* confidence interval of the population mean based on a sample of *n = 25* with *s = 0.05*, the critical value of *t* is
   A) *2.492*
   B) *2.797*
   C) *2.787*
   D) *2.485*

<u>Solution</u>: t = t(df = n − 1, c = 0.99) = t(df = 24, c = 0.99) = 2.797*

Use the following to answer questions 10-11:
Twelve runners are asked to run a 10-kilometer race on each of two consecutive weeks. In one of the races, the runners wear one brand of shoe and in the other a different brand. All runners are timed and are asked to run their best in each race. The results (in minutes) are given below:

Runner	1	2	3	4	5	6	7	8	9	10	11	12
Brand 1	31.23	29.33	30.5	32.2	33.08	31.52	30.68	31.05	33	29.67	30.55	32.12
Brand 2	32.02	28.98	30.63	32.67	32.95	31.53	30.83	31.1	33.12	29.5	30.57	32.2

Use the 10% significance level for matched pairs to determine if there is evidence that times using Brand 1 tend to be faster than times using Brand 2. The Minitab output is listed below:

```
0.00000   hypothesized value
31.24417  mean Brand 1
31.34167  mean Brand 2
-0.09750  mean difference  (Brand 1 - Brand 2)
0.29579   std. dev.
0.08539   std. error
12        n
11        df
```

```
-1.14     t
.1389     p-value (one-tailed, lower)

-0.25085  confidence interval 90.% lower
0.05585   confidence interval 90.% upper
0.15335   margin of error
```

10. The formula for the test statistic is:

A) $t = \dfrac{\bar{x} - \mu}{s/\sqrt{n}}$

B) $t = \dfrac{\bar{d}}{s_d/\sqrt{n}}$

C) $z = \dfrac{\bar{x}_1 - \bar{x}_2}{\sqrt{\dfrac{(\sigma_1)^2}{n_1} + \dfrac{(\sigma_2)^2}{n_2}}}$

D) $t = \dfrac{\bar{x}_1 - \bar{x}_2}{\sqrt{\dfrac{(s_1)^2}{n_1} + \dfrac{(s_2)^2}{n_2}}}$

<u>Solution:</u> since we compare two different results of each runner, we perform a paired t-test.

11. It can be concluded at the 10% significance level that

A) since H_0 is rejected, the Brand 1 is faster than the Brand 2.
B) since H_0 is rejected, the Brand 1 is not faster than the Brand 2.
C) since H_0 is not rejected, the Brand 1 is faster than the Brand 2.
D) since H_0 is not rejected, the Brand 1 is not faster than the Brand 2.

<u>Solution:</u> since p-value = 0.1389 is greater than the level of significance $\alpha = 0.1$, the null hypothesis H_0 is not rejected. Therefore, the Brand 1 is not faster than the Brand 2.

<u>Use the following to answer questions 12-13:</u>
Are avid readers more likely to wear glasses than those who read less frequently? Three hundred men in Ohio were selected at random and characterized as to whether they wore glasses and whether the amount of reading they did was above average, average, or below average. The results and Minitab output are presented below:

	Glasses?	
Amount of reading	Yes	No
Above average	45	30
Average	45	70
Below average	30	50
Total	120	150

Chi-square Contingency Table Test for Independence

		Yes	No	Total
Above average	Observed	45	30	**75**
	Expected	***	41.67	**75.00**
Average	Observed	45	70	115
	Expected	51.11	63.89	115.00
Below average	Observed	30	50	80
	Expected	35.56	44.44	80.00
Total	Observed	**120**	150	**270**
	Expected	**120.00**	150.00	**270.00**

10.23	chi-square
2	df
.0060	p-value

12. Suppose we wish to test the null hypothesis that there is no association between the amount of reading and wearing glasses. Under the null hypothesis, what is the expected number of above average readers who wear glasses?
    A) *45*
    B) *33.33*
    C) *75*
    D) *120*

<u>Solution</u>: $E = \frac{(row\,total)\times(column\ total)}{total} = \frac{120\times 75}{270} = 33.33$

13. It can be concluded at the *1%* significance level that
    A) there is no association between the amount of reading and wearing glasses.
    B) an association between the amount of reading and wearing glasses exists.
    C) there is insufficient information to make any conclusions.
    D) there is no association between the amount of average reading and wearing glasses.

<u>Solution</u>: since p-value = 0.006 is less than the level of significance α = 0.01, the null hypothesis H_0 is rejected. Therefore, an association between the amount of reading and wearing glasses exists.

Use the following to answer questions 14-15:
A U.S. newspaper is conducting a statewide survey concerning the race for governor. The newspaper takes a simple random sample of *100* registered voters and determines that *53* of them will vote for the Democratic candidate. Is there evidence that a majority of the population will vote for the Democratic candidate? To answer this, they will test the hypotheses H_0: $p = 0.5$ versus H_a: $p > 0.5$ at the *5%* significance level.

14. What is the test statistic?
    A) $z = 0.600$
    B) $z = 0.601$
    C) $p = 0.530$
    D) $p = 0.270$

Solution: this is the right-tailed test of the population proportion: $z = \frac{\hat{p} - p}{\sqrt{\frac{p(1-p)}{n}}} = \frac{53/100 - 0.5}{\sqrt{\frac{0.5(1-0.5)}{100}}} = 0.6$

15. If the *p-value = 0.27*, which of the following statements is true?
    A) H_0 is rejected at the *5%* significance level, so a majority of the population will vote for the Democratic candidate.
    B) H_0 is not rejected at the *5%* significance level, so a majority of the population will vote for the Democratic candidate.
    C) H_0 is rejected at the *5%* significance level, so a majority of the population will not vote for the Democratic candidate.
    D) H_0 is not rejected at the *5%* significance level, so a majority of the population will not vote for the Democratic candidate.

Solution: since p-value = 0.27 is greater than the level of significance $\alpha = 0.05$, the null hypothesis H_0 is not rejected. Therefore, a majority of the population will not vote for the Democratic candidate.

---

16. Which of the following statements is true?
    A) The probability of making a Type II error and the level of significance are the same.
    B) The probability of making a Type II error and the power of the test are the same.
    C) The probability of making a Type I error and the power of the test are the same.
    D) The power of a test decreases as the level of significance decreases.

Solution: Decreasing (increasing) the probability of the type I error (α), we increase (decrease) the probability of the type II error (β). The power of the test is equal to $1 - \beta$, so it decreases (increases) as the level of significance decreases (increases).