# Answers

## Week 1

Theory

a. Population mean: $\bar{x}$, Sample mean, $\mu$
b. $H_0$ is what we expect to see happen based on the knowledge we have, $H_a$ is what we suppose happens if the null hypothesis is not true
c. The p-value is the probability that you find what you find
d. The limit after which we state that the null hypothesis is not true and we assume the alternative hypothesis as probable.

Application

2 and 5 are correct

## Week 2

a. Event A alters event B or vice versa.
b. Event A and B cannot exist or occur at the same time.
c. They cannot.
d. **1.** P(A/B) = P(A) + P(B) – P(A&B) **2.** P(A/B) = P(A) + P(B) **3.** 1 – P(A) = P(B) **4.** P(A&B) = P(A) × P(B|A) **5.** P(A&B) = P(A) × P(B)
e. There is no chance of you knowing the outcome
f. P(x) = F(x)/n
g. $\mu_{\bar{x}} = \sum \bar{x}_n \times p_n$
h. $\sigma^2_x = \sum(x_i - \mu_x)^2 \times p_i$
i. $\mu_{x+y} = \mu_x + \mu_y$
j. $\sigma^2_{x+y} = \sigma^2_x + \sigma^2_y + 2p_{xy} \times \sigma_x \times \sigma_y$

## Week 3

a. Numerical data can only take certain, definite values, while categorical values can be anything.
b. When they don't alter or influence each other's results in any way.
c. Goodness of fit, Independence and homogeneity
d. Number of cells that can be filled in freely with only restrictions of marginal totals
e. Number of cells that can be filled in freely with only restrictions of marginal totals – 1
f. $fe$ (A and B) = $f$(A) × $f$(B) / n
g. $X^2 = \sum$ (Observed – Expected) / Expected

## Week 4

a. $H_0$ is regular
b. $S_x$
c. $\sigma_x$
d. T = $\bar{x}$ - $\mu$ / $S_x$

e. $z = \bar{x} - \mu / \sigma_x$
f. One-sided

## Week 5

a. $H_0: \mu_d = 0$ and ( if one-sided) $H_a: \mu_d < / > 0$ or (if two- sided) $H_a: \mu_d \neq 0$
b. $H_0: \mu_1 - \mu_2 = 0$ and ( if one-sided) $H_a: \mu_1 - \mu_2 < / > 0$ or (if two- sided) $H_a: \mu_1 - \mu_2 \neq 0$
c. 1. Df = number of pairs observed – 1. 2. Df = The smallest of the two n's – 1. 3. $n_1 + n_2 - 2$.
d. $t = \bar{x}_1 - \bar{x}_2 / se$

$$se_{\bar{x}_1 - \bar{x}_2} = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

e. Because most of the time $\mu_1 - \mu_2$ is equal to 0

f. $t = D/se$ $se_d = S_d / \sqrt{n}$

## Week 6

Theory

a. That in 95% of all samples, the mean falls within two standard errors of the population mean.
b. $CI_{1-\alpha} = \bar{x} \pm t^* \times SE$
   a. $t^* = t_{\alpha/2}(df)$
   b. $SE = s/\sqrt{n}$
c. $CI_{1-\alpha} = (\bar{x}_1 - \bar{x}_2) \pm t^* \times SE_{\bar{x}1 - \bar{x}2}$
d. Effect size: $\eta^2 = t^2 / (t^2 + df)$
e. Cohen's d: $d = \bar{x} - \mu_0 / s$

Application

D is correct, a is technically true as well but has nothing to do with the interpretation of the 95% CI

## Week 7

Theory

a. Distribution- free data that does not, or rarely makes claims about a population, as well as for resampling tests or ordinal data
b. When dealing with ordinal data or with continuous data that has a small n and a skewed sample distribution, in the case when you don't know anything about the population.
c. Paired = signed rank, Independent = rank sum
d. You take the average.

e. $\mu = n(n+1) / 4$ $\sigma = n(n+1)(2n+1)/24$
f. $\mu = n_1 (n_1 + n_2 + 1) / 2$ $\sigma = n_1 \times n_2 (n_1 + n_2 + 1)/12$
g. $z = T + - \mu_{T^*} / \sigma_{T^*}$ $T = W - \mu_W / \sigma_W$

Application

a. 4, 4, 5, 6, 6, **7, 8,** 9, 9, 9, 10, **14, 14, 15, 15, 17, 18, 19, 22, 36**

| | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 4 | 4 | 5 | 6 | 6 | **7** | **8** | 9 | 9 | 9 | 10 | **14** | **14** | **15** | **15** | **17** | **18** | **19** | **22** | **36** |
| 1.5 | 1.5 | 3 | 4.5 | 4.5 | **6** | **7** | 9 | 9 | 9 | 11 | **12.5** | **12,5** | **14.5** | **14.5** | **16** | **17** | **18** | **19** | **20** |

No brain damage (n=9) = 53 Mean is 115,5
**brain damage (n=11) = 157 Mean is 94.5**
Standard dev. = 13.16
Z = -3.153
P = 0.016
$H_0$ = rejected

## Week 8

a. $\varphi = \sqrt{(x^2/n)}$

b. $V = \sqrt{(x^2/n(k-1))}$