Practice Questions for Exam 3

Human beta-endorphin (HBE) is a hormone secreted by the pituitary gland under condition of stress. An exercise physiologist measured the resting (un-stresses) blood concentration of HBE in three groups of men: i) who had just entered a physical fitness program ii) who had been jogging regularly for some time and iii) sedentary people.

The HBE level (pg/ml) is shown in the following table:

Entrants	38.7	41.2	39.3	37.4	38.4	36.7	41.2	43.6	37.2		
Joggers	35.7	31.2	33.4	38.9	35.2	35.1	34.3	36.2			
Sedentary	42.5	43.8	40.3	38.7	41.2	45.6	44.3	36.8	43.2	37.4	38.8

1. The appropriate null hypothesis to test if HBE level is affected by level exercise/fitness.
   a) $H_0: p_1 = p_2 = p_3 \quad vs. \quad H_A: not\ H_0$
   b) $H_0: \mu_1 = \mu_2 = \mu_3 \quad vs. \quad H_A: not\ H_0$
   c) $H_0: proportion\ of\ high\ HBE\ same\ in\ three\ groups \quad vs. \quad H_A: not\ H_0$
   d) None of the above.

Following is the Rcode and output for the analysis of the data.

```
HBE=c(38.7,41.2,39.3,37.4,38.4,36.7,41.2,43.6,37.2,35.7,31.2,33.4,38.9,35.2
,35.1,34.3,36.2,42.5,43.8,40.3,38.7,41.2,45.6,44.3,36.8,
43.2,37.4,38.8)
groups=c(1,1,1,1,1,1,1,1,1,2,2,2,2,2,2,2,2,3,3,3,3,3,3,3,3,3,3,3)
groups=factor(groups)
fit=aov(HBE~groups)
summary(fit)

                Df    Sum Sq    Mean Sq     F value    Pr(>F)
groups           2    178.3       89.14       13.56     0.000103 ***
Residuals    25    164.3       6.57
---
```

2. The categorical variable in this study is
   a) Exercise groups b) HBE level c) Both a) and b) d) None of a) or b)

3. The value of the F-test statistic is
   a) 89.14 b) 0.000103 c) 13.56 d) 6.57

4. The p-value for the test is:
   a) 89.14 b) 0.000103 c) 13.56 d) 6.57

5. At 5% level of significance we conclude that:
   a) The three groups have significantly different average HBE level.
   b) The three groups do not have significantly different average HBE level.
   c) Joggers have significantly higher HBE level than other groups.
   d) Sedentary people have significantly lower HBE level than other groups.

Migraine headache patients took part in a double-blind clinical trial to assess experimental surgery. 49 patients were assigned to a real surgery and 26 were assigned to a sham surgery. Table below shows the outcome:

	Real	Sham	**Total**
Success	41	15	**56**
Failure	8	11	**19**
Total	**49**	**26**	**75**

6. The estimated odds ratio for success for real to sham group is:
   a) 1.45 b) 3.76 c) 2.34 d) 4.11

The following show R code and output from the data

```
> migraine=matrix(c(41,8,15,11),nrow=2)
> fisher.test(migraine)

        Fisher's Exact Test for Count Data
data:  migraine
p-value = 0.02409
alternative hypothesis: true odds ratio is
not equal to 1
95 percent confidence interval:
  1.109644    12.884047
```

```
> migraine=matrix(c(41,8,15,11),nrow=2)
> prop.test(migraine)

        2-sample test for equality of
proportions with continuity correction


data:  migraine

X-squared = 4.7661, df = 1, p-value = 0.02902

alternative hypothesis: two.sided

95 percent confidence interval:

 0.02537031 0.59681014
```

Let us assume p_1 and p_2 denote the chances of having success outcomes while in real surgery and sham surgery respectively and θ denotes the odds ratio of having success outcomes from real surgery to sham surgery.

7. An appropriate test to see if chance of a success while in real surgery is any different than chance of a success while in sham surgery, will be:
   a) $H_0: p_1 = p_2 \;\; vs. \;\; H_A: p_1 \neq p_2$
   b) $H_0: p_1 - p_2 = 0 \;\; vs. \;\; H_A: p_1 - p_2 \neq 0$
   c) $H_0: \theta = 1 \;\; vs. \;\; H_A: \theta \neq 1$
   d) All of the above.

8. What's a 95% confidence interval for $p_1 - p_2$?

9. In R, the code "prop.test" actually tests the hypotheses:
   a) $H_0: p_1 - p_2 = 0 \;\; vs. \;\; H_A: p_1 - p_2 \neq 0$
   b) $H_0: \theta = 1 \;\; vs. \;\; H_A: \theta \neq 1$
   c) $H_0: p_1/p_2 = 1 \;\; vs. \;\; H_A: p_1/p_2 \neq 1$
   d) $H_0: \mu_1 - \mu_2 = 0 \;\; vs. \;\; H_A: \mu_1 - \mu_2 \neq 0$

10. In R, the code "fisher.test" actually tests the hypotheses:
   a) $H_0: p_1 - p_2 = 0 \;\; vs. \;\; H_A: p_1 - p_2 \neq 0$
   b) $H_0: \theta = 1 \;\; vs. \;\; H_A: \theta \neq 1$
   c) $H_0: p_1/p_2 = 1 \;\; vs. \;\; H_A: p_1/p_2 \neq 1$
   d) $H_0: \mu_1 - \mu_2 = 0 \;\; vs. \; H_A: \mu_1 - \mu_2 \neq 0$

11. What's a 95% confidence interval for the population odds ratio θ?

12. The appropriate test to see if the surgery work is:
   a) Test for independence
   b) goodness of fit test
   c) Simpson's paradox
   d) Analysis of Variance test

13. The conclusion from the fisher test is:
   a) Success or failure is not associated with the fact the surgery is real or sham.
   b) Success or failure is associated with the fact the surgery is real or sham.
   c) Patients have better odds of having the real surgery than the sham surgery.
   d) Patients have lesser odds of having the real surgery than the sham surgery.

14. Suppose the study was done in two clinics, A and B.

Clinic A:

	Real	Sham
Success	31	9
Failure	2	2

Clinic B:

	Real	Sham
Success	11	6
Failure	6	9

```
migraineA=matrix(c(31,2,9,2),nrow=2
)
> fisher.test(migraineA)

        Fisher's Exact Test for
Count Data
data:  migraineA
p-value = 0. 2565
alternative hypothesis: true odds
ratio is not equal to 1

95 percent confidence interval:
  0.2135324 52.0460176

sample estimates:

odds ratio

  3.327504
```

```
migraineB=matrix(c(11,6,6,9),nrow=2
)
> fisher.test(migraineB)

        Fisher's Exact Test for
Count Data
data:  migraineB
p-value = 0.2872alternative
hypothesis: true odds ratio is not
equal to 1

95 percent confidence interval:

  0.5335528 14.6430187

sample estimates:

odds ratio

  2.660669
```

After stratifying clinic, what's the odds ratio within each stratified level and its p value?

Is the association still significant? (Simpson's paradox)

Cochran-Mantel-Haenszel test

```
> M <- array(c(31,2,9,2,11,6,6,9),dim=c(2,2,2))

> mantelhaen.test(M)

        Mantel-Haenszel chi-squared test with continuity correction

data:  M

Mantel-Haenszel X-squared = 2.2194, df = 1, p-value = 0.1363

alternative hypothesis: true common odds ratio is not equal to 1

95 percent confidence interval:
```

0.8943295 9.6332641

sample estimates:

common odds ratio

2.935185

What does Cochran-Mantel-Haenszel test test for?

In a breeding experiment, white chickens with small combs were mated and produced 190 offspring of the types shown in below table. Are these data consistent with the Mendelian expected ratios of 9:3:3:1 for the four types?

Type	**Number of offspring**
White feathers, small comb	111
White feathers, large comb	37
Dark feathers, small comb	34
Dark feathers, large comb	8
Total	190

15. Let p1, p2, p3 and p4 be proportions of different type offspring. To verify if the data are consistent with the Mendelian expected ratios of 9:3:3:1 for the four types, we should test the following hypothesis
    $a)$: $p_1 = p_2 = p_3 = p_4$ $vs.$ H_A: H_0 is false.
    b) H_0: $\theta = 1$ $vs.$ H_A: $\theta \neq 1$
    c) H_0: $p_1 = \frac{9}{16}$; $p_2 = \frac{3}{16}$; $p_3 = \frac{3}{16}$; $p_4 = \frac{1}{16}$ $vs.$ H_A: H_0 is false.
    d) None of the above.

16. What's the expected number of offspring for the type of Dark feathers and small comb?
    a) 34 b) $\frac{1}{4} * 34$ c) $\frac{3}{16} * 190$ d) $\frac{1}{3} * 34$

17. Appropriate test to use in this case:
    a) Test for independence
    b) Goodness of fit test
    c) Simpson's paradox
    d) Fisher's test.

13. What’s the degrees of freedom of the test?

14. The Rcode to be used for the hypothesis test is:

a)
```
type=c(11,37,34,8)
chisq.test(type)
```

b)
```
type=c(11,37,34,8)
prob=c(0.5625, 0.1875, 0.1875, 0.0625)
chisq.test(type,p=prob)
```

c)
```
type=c(11,37,34,8)
fisher.test(color)
```

d)
```
type=c(11,37,34,8)
prob=c(0.5625, 0.1875, 0.1875, 0.0625)
simpson.test(color)
```

15. Answer each statement below True or False

____ In 2x2 tables, statistical independence is equivalent to a population odds ratio value of $\theta = 1.0$.

____Fisher’s exact test is a test of the null hypothesis of independence for 2×2 contingency tables that fixes the row and column totals and uses a hypergeometric distribution for the count in the first cell. For a one-sided alternative of a positive association (i.e., odds ratio > 1), the P-value is the sum of the probabilities of all those tables that have count in the first cell at least as large as observed, for the given marginal totals.

____ Fisher’s exact test for a contingency table is suitable when expected frequencies are small.

____ The difference of proportions, relative risk, and odds ratio are valid measures for summarizing 2×2 tables for either prospective or retrospective (e.g., case-control) studies.

____ The default chi-squared test gives larger p-value than the chi-squared test with Yates’ continuity correction.

____ When an odds ratio is calculated from a 2x2 table, the odds ratio will not be changed of the order of the rows and order of the columns is reversed.

____ The degrees of freedom for a chi-square test of association using 3*5 table are 8.

16. To investigate the dependence of energy expenditure on body build, researchers used underwater weighing techniques to determine the fat-free body mass for each of seven men. They also measured the total 24-hour energy expenditure for each man during conditions of quite sedentary activity. The results are shown below:

Subject	Fat-free mass X (kg)	Energy expenditure Y (kcal)
1	49.3	1894
2	59.3	2050
3	68.3	2353
4	48.1	1838
5	57.6	1948
6	78.1	2528
7	76.1	2568

Here's the R output for this data:

```
> fit=lm(energy~mass)
> summary(fit)
Call:
lm(formula = energy ~ mass)

Residuals:
    1      2      3      4      5      6      7
  53.22  -40.89   37.00   27.24 -100.37  -33.11   56.91

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  607.703    138.765   4.379   0.00716 **
mass          25.012      2.189  11.427 8.99e-05 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 64.85 on 5 degrees of freedom
Multiple R-squared:  0.9631,    Adjusted R-squared:  0.9557
F-statistic: 130.6 on 1 and 5 DF,  p-value: 8.988e-05

The fitted line is:  energy = 607.7 + 25.0 (mass)
> confint(fit)
                 2.5 %       97.5 %
```

(Intercept)	250.99778	964.40909
mass	19.38506	30.63818

17. What's your regression equation?

18. Plugging 60 into the regression equation would yield
a) An estimate of the mean energy expenditure
b) An estimate of the mean fat free mass
c) An estimate of the variation in energy expenditure
d) An estimate of the variation of fat free mass

19. The coefficient of determination (r^2) is
a) 96.31% of the variation in energy expenditure is explained by fat free mass
b) 96.31% of the variation in fat free mass is explained by energy expenditure
c) Predictions of energy expenditure tend to be off by 0.9631 in general
d) energy expenditure goes up by 0.9631 for every 1 kg fat free mass change

20. What's 95% confidence interval for β_1

21. The correlation between energy expenditure and fat-free mass?
a) 0.9631 b) 0.9557 c) 0.9814 d) can't tell from the output