JARAMOGI OGINGA ODINGA UNIVERSITY OF SCIENCE AND TECHNOLOGY
SCHOOL OF HEALTH SCIENCES

UNIVERSITY EXAMINATION FOR DEGREE OF MASTER PUBLIC HEALTH

1ST YEAR 2ND SEMESTER 2022/2023 ACADEMIC YEAR

KISUMU CAMPUS

COURSE CODE: HES 5123

COURSE TITLE: ADVANCED BIOSTATISTICS

EXAM VENUE:	**STREAM: (MPH)**
DATE:	**EXAM SESSION:**
TIME:	**3.00 HOURS**

Instructions:

1. Answer **Question ONE (compulsory)** and any other THREE questions.
2. Candidates MUST not to write anything on the question paper.
3. Candidates MUST hand in their answer booklets to the invigilator while in the examination room.

SECTION A

Answer question one(Compulsary)

1. **Question one (10 marks).**

a) Show that $se(\hat{\pi}) = \frac{1}{\sqrt{-l''(\hat{\pi})}}$ where $se(\hat{\pi})$ is the standard error of the estimated population prevalence and $l''(\hat{\pi})$ is the second derivative of the log-likelihood function of the estimated population prevalence. (3 marks)

b) Show invariance of Maximum likelihood Estimator $If\ \theta = f(\pi), then\ \hat{\theta} = f(\hat{\pi})$. (2 marks)

c) Derive the Wool's formulae for log odds $se[\log(\widehat{\Omega})] = \sqrt{\frac{1}{x} + \frac{1}{n-x}}$ for X being the number of individuals with disease, $se[\log(\widehat{\Omega})]$ is the standard error of the log-odds and n is the sample size. (3 marks)

d) Hence similarly for (c) above show that $se[\log(\widehat{\pi})] = \sqrt{\frac{1-\widehat{\pi}}{n\widehat{\pi}}}$. (2 marks)

SECTION B

Answer any three Questions

2. Question two (20 marks).

In a randomized trial patients infected by helicobacter pylori were randomly allocated to treatment by drug combination A or treatment by drug combination B. At the end of the study, the non-cure rates are to be compared between the two groups, using the risk difference or the risk ratio as effect measure.

RESIST resistant against one of the drugs in the combination

0 = no, 1 = yes

CURE cured : 1 = not cured, 0 = cured

TREAT treatment : 0 = drug combination A , 1 = drug combination B

Treatment * Cured * Resistant against one of the drugs in the combination Crosstabulation

Count

Resistant against one of the drugs in the combination			Cured: cured	Cured: not cured	Total
no	Treatment	drug combination A	111	3	114
		drug combination B	99	6	105
	Total		**210**	**9**	**219**
yes	Treatment	drug combination A	90	9	99
		drug combination B	75	12	87
	Total		**165**	**21**	**186**

Some SAS output

```
                    Table of TREAT by CURE

        TREAT(Treatment)      CURE(Cured)

        Frequency        ,
        Row Pct          ,cured    ,not cure,  Total
```

		d	
drug combination A	201 94.37	12 5.63	213
drug combination B	174 90.63	18 9.38	192
Total	375	30	405

Statistics for Table of TREAT by CURE

Column 2 Risk Estimates

	Risk	ASE	(Asymptotic) 95% Confidence Limits		(Exact) 95% Confidence Limits	
Row 1	0.0563	0.0158	0.0254	0.0873	0.0294	0.0963
Row 2	0.0938	0.0210	0.0525	0.1350	0.0565	0.1441
Total	0.0741	0.0130	0.0486	0.0996	0.0505	0.1041
Difference	-0.0374	0.0263	-0.0890	0.0142		

Difference is (Row 1 - Row 2)

Statistics for Table of TREAT by CURE

Estimates of the Relative Risk (Row1/Row2)

Type of Study	Value	95% Confidence Limits	
Case-Control (Odds Ratio)	1.7328	0.8119	3.6981
Cohort (Col1 Risk)	1.0413	0.9845	1.1014
Cohort (Col2 Risk)	0.6009	0.2973	1.2149

Sample Size = 405

Model Information

Data Set	WORK.HELIPYL	
Distribution	Binomial	
Link Function	Identity	
Dependent Variable	CURE	Cured

Number of Observations Read	405
Number of Observations Used	405
Number of Events	30
Number of Trials	405

Response Profile

Ordered Value	CURE	Total Frequency
1	not cured	30
2	cured	375

PROC GENMOD is modeling the probability that CURE='not cured'.

Criteria For Assessing Goodness Of Fit

Criterion	DF	Value	Value/DF
Log Likelihood		-105.9089	
Full Log Likelihood		-105.9089	
AIC (smaller is better)		215.8177	
AICC (smaller is better)		215.8476	
BIC (smaller is better)		223.8255	

Algorithm converged.

Analysis Of Maximum Likelihood Parameter Estimates

Parameter	DF	Estimate	Standard Error	Wald 95% Confidence Limits		Likelihood Ratio 95% Confidence Limits		Wald Chi-Square	Pr > ChiSq
Intercept	1	0.0563	0.0158	0.0254	0.0873	0.0306	0.0927	12.72	0.0004
TREAT	1	0.0374	0.0263	-0.0142	0.0890	-0.0137	0.0914	2.02	0.1550
Scale	0	1.0000	0.0000	1.0000	1.0000	1.0000	1.0000		

NOTE: The scale parameter was held fixed.

Relative risk model

The GENMOD Procedure

Model Information

Data Set	WORK.HELIPYL	
Distribution	Binomial	
Link Function	Log	
Dependent Variable	CURE	Cured

Number of Observations Read	405
Number of Observations Used	405
Number of Events	30
Number of Trials	405

Response Profile

Ordered Value	CURE	Total Frequency
1	not cured	30
2	cured	375

PROC GENMOD is modeling the probability that CURE='not cured'.

Parameter Information

Parameter	Effect
Prm1	Intercept
Prm2	TREAT

Criteria For Assessing Goodness Of Fit

Criterion	DF	Value	Value/DF
Log Likelihood		-105.9089	
Full Log Likelihood		-105.9089	
AIC (smaller is better)		215.8177	
AICC (smaller is better)		215.8476	
BIC (smaller is better)		223.8255	

Algorithm converged.

Analysis Of Maximum Likelihood Parameter Estimates

Parameter	DF	Estimate	Standard Error	Wald 95% Confidence Limits		Likelihood Ratio 95% Confidence Limits		Wald Chi-Square	Pr > ChiSq
Intercept	1	-2.8764	0.2804	-3.4260	-2.3268	-3.4876	-2.3786	105.21	<.0001
TREAT	1	0.5093	0.3591	-0.1947	1.2132	-0.1844	1.2416	2.01	0.1562
Scale	0	1.0000	0.0000	1.0000	1.0000	1.0000	1.0000		

NOTE: The scale parameter was held fixed.

Contrast Estimate Results

Label	Mean Estimate	Mean Confidence Limits		L'Beta Estimate	Standard Error	Alpha	L'Beta Confidence Limits	
TreatmentB	1.6641	0.8231	3.3641	0.5093	0.3591	0.05	-0.1947	1.2132
Exp(TreatmentB)				1.6641	0.5976	0.05	0.8231	3.3641

Contrast Estimate Results

Label	Chi-Square	Pr > ChiSq
TreatmentB	2.01	0.1562
Exp(TreatmentB)		

Relative risk model

The FREQ Procedure

Table 1 of TREAT by CURE
Controlling for RESIST=no

TREAT(Treatment)   CURE(Cured)

Frequency / Row Pct	cured	not cured	Total
drug combination A	111 97.37	3 2.63	114
drug combination B	99 94.29	6 5.71	105
Total	210	9	219

Statistics for Table 1 of TREAT by CURE
Controlling for RESIST=no

Column 1 Risk Estimates

	Risk	ASE	(Asymptotic) 95% Confidence Limits		(Exact) 95% Confidence Limits	
Row 1	0.9737	0.0150	0.9443	1.0000	0.9250	0.9945
Row 2	0.9429	0.0227	0.8985	0.9873	0.8798	0.9787
Total	0.9589	0.0134	0.9326	0.9852	0.9234	0.9810
Difference	0.0308	0.0272	-0.0224	0.0841		

Difference is (Row 1 - Row 2)

Column 2 Risk Estimates

	Risk	ASE	(Asymptotic) 95% Confidence Limits		(Exact) 95% Confidence Limits	
Row 1	0.0263	0.0150	0.0000	0.0557	0.0055	0.0750
Row 2	0.0571	0.0227	0.0127	0.1015	0.0213	0.1202
Total	0.0411	0.0134	0.0148	0.0674	0.0190	0.0766
Difference	-0.0308	0.0272	-0.0841	0.0224		

Difference is (Row 1 - Row 2)

Relative risk model

The FREQ Procedure

Statistics for Table 1 of TREAT by CURE
Controlling for RESIST=no

Estimates of the Relative Risk (Row1/Row2)

Type of Study	Value	95% Confidence Limits	
Case-Control (Odds Ratio)	2.2424	0.5463	9.2045
Cohort (Col1 Risk)	1.0327	0.9765	1.0921

Type of Study	Value	95% Confidence Limits	
Cohort (Col2 Risk)	0.4605	0.1182	1.7949

Sample Size = 219

Table 2 of TREAT by CURE
Controlling for RESIST=yes

TREAT(Treatment) Frequency / Row Pct	CURE(Cured): cured	not cured	Total
drug combination A	90 90.91	9 9.09	99
drug combination B	75 86.21	12 13.79	87
Total	165	21	186

Statistics for Table 2 of TREAT by CURE
Controlling for RESIST=yes

Column 1 Risk Estimates

	Risk	ASE	(Asymptotic) 95% Confidence Limits		(Exact) 95% Confidence Limits	
Row 1	0.9091	0.0289	0.8525	0.9657	0.8344	0.9576
Row 2	0.8621	0.0370	0.7896	0.9345	0.7715	0.9266
Total	0.8871	0.0232	0.8416	0.9326	0.8326	0.9287
Difference	0.0470	0.0469	-0.0449	0.1390		

Difference is (Row 1 - Row 2)

Relative risk model

The FREQ Procedure

Statistics for Table 2 of TREAT by CURE
Controlling for RESIST=yes

Column 2 Risk Estimates

	Risk	ASE	(Asymptotic) 95% Confidence Limits		(Exact) 95% Confidence Limits	
Row 1	0.0909	0.0289	0.0343	0.1475	0.0424	0.1656
Row 2	0.1379	0.0370	0.0655	0.2104	0.0734	0.2285
Total	0.1129	0.0232	0.0674	0.1584	0.0713	0.1674
Difference	-0.0470	0.0469	-0.1390	0.0449		

Difference is (Row 1 - Row 2)

Estimates of the Relative Risk (Row1/Row2)

Type of Study	Value	95% Confidence Limits	
Case-Control (Odds Ratio)	1.6000	0.6396	4.0028
Cohort (Col1 Risk)	1.0545	0.9498	1.1708
Cohort (Col2 Risk)	0.6591	0.2918	1.4888

Sample Size = 186

Answer the following questions

a. Compute the risk difference and a relative risk. (2 marks)
b. Compare the risk of not getting cured with either treatment A or B. Find the risk difference/risk ratio and its confidence interval for the non-cure rate. (hint transformation cure=1). (2 marks)

c. Compute the Wald's 95%CI. (2 marks)
d. What is the interpretation of the estimated regression coefficients? (2 marks)
e. Some of the patients are resistant to one of the drugs in the drug combination, others are not resistant. Adjust the estimates of the risk difference for resistance. (2 marks)
f. Compute the risk differences in both strata. (4 marks)
g. Calculate the weighted mean of the two risk difference using the weight factor for each stratum: One over the squared standard error. (4 marks)
h. Is there evidence of heterogeneity for both RD and RR? (Hint: use interaction term that tests that RD or RR in two strata is equal) (2 marks)

3. Question three (20 marks).

1. Presence of a certain element of the set of teeth in babies, depending on age

Y=1/0 if element present/absent

X=age at examination (weeks)

Using binary logistic regression in SPSS gives the following:

Y

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0	38	76.0	76.0	76.0
	1	12	24.0	24.0	100.0
	Total	50	100.0	100.0	

Block 0

Variables in the Equation

		B	S.E.	Wald	Df	Sig.	Exp(B)
Step 0	Constant	-1.153	.331	12.117	1	.000	.316

Iteration History(a,b,c,d)

Iteration		-2 Log likelihood	Coefficients	
			Constant	X
Step 1	1	36.215	-3.827	.095
	2	29.677	-6.483	.162
	3	27.743	-8.796	.220
	4	27.474	-10.043	.251
	5	27.467	-10.287	.257

6	27.467	-10.295	.257
7	27.467	-10.295	.257

a Method: Forward Stepwise (Wald)
b Constant is included in the model.
c Initial -2 Log Likelihood: 55.108
d Estimation terminated at iteration number 7 because parameter estimates changed by less than .001.

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)	95.0% C.I.for EXP(B)	
								Lower	Upper
Step 1(a)	X	.257	.078	10.727	1	.001	1.293	1.109	1.508
	Constant	-10.295	3.066	11.275	1	.001	.000		

a Variable(s) entered on step 1: X.

Correlation Matrix

		Constant	X
Step 1	Constant	1.000	-.987
	X	-.987	1.000

a. Estimate of the covariance matrix, hence what are the standard errors(s_0) and (s_1)? (4 marks)
b. What is the correlation between $\hat{\beta}_0$ and $\hat{\beta}_1$. (2 marks)
c. Give the 95%CI for β_1 using the Wald's method. (2 marks)
d. What is the probability that a 40 week old will have the element?. (8 marks)
e. Test for $H_0: \beta_1 = 0$ with three methods (follow SPSS output). (4 marks)

4. Question four (20 marks).

2. In a random sample from the population of a rural area in a certain developing country the following variables, among others, were observed on 328 persons.

SYS systolic blood pressure (mmHg)

PULSE pulse rate (beats/min)

SES social economic status (1=lower class, 2=middle class, 3=upper class)

This problem concentrates on the differences in mean systolic blood pressure between the three social economic classes corrected for pulse frequency. Three multiple regression models were filled using SPSS. Part of the output is given below.

Model 1:

Variables Entered/Removed(b)

Model	Variables Entered	Variables Removed	Method
1	middle social economic class, low social economic status(a)	.	Enter

a All requested variables entered.
b Dependent Variable: systolic blood pressure (mmHg)

ANOVA(b)

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	4019.437	2	2009.719	6.898	.001(a)
	Residual	94683.840	325	291.335		
	Total	98703.277	327			

a Predictors: (Constant), middle social economic class, low social economic status
b Dependent Variable: systolic blood pressure (mmHg)

Coefficients(a)

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	126.381	1.002		126.175	.000
	low social economic status	-2.196	1.307	-.095	-1.681	.094
	middle social economic class	-3.645	1.330	-.155	-2.741	.006

a Dependent Variable: systolic blood pressure (mmHg)

Model 2:

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.169(a)	.029	.026	17.15045
2	.258(b)	.067	.058	16.86281
3	.259(c)	.067	.055	16.88669

a Predictors: (Constant), pulse frequency (beats/min)
b Predictors: (Constant), pulse frequency (beats/min), low social economic status, middle social economic class
c Predictors: (Constant), pulse frequency (beats/min), low social economic status, middle social economic class, squared pulse rate

ANOVA(d)

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	2814.288	1	2814.288	9.568	.002(a)
	Residual	95888.989	326	294.138		
	Total	98703.277	327			
2	Regression	6572.496	3	2190.832	7.705	.000(b)
	Residual	92130.781	324	284.354		
	Total	98703.277	327			
3	Regression	6596.497	4	1649.124	5.783	.000(c)
	Residual	92106.780	323	285.160		
	Total	98703.277	327			

a Predictors: (Constant), pulse frequency (beats/min)
b Predictors: (Constant), pulse frequency (beats/min), low social economic status, middle social economic class
c Predictors: (Constant), pulse frequency (beats/min), low social economic status, middle social economic class, squared pulse rate
d Dependent Variable: systolic blood pressure (mmHg)

Coefficients(a)

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	104.616	6.711		15.588	.000
	pulse frequency (beats/min)	.250	.081	.169	3.093	.002
2	(Constant)	106.752	6.625		16.113	.000
	pulse frequency (beats/min)	.239	.080	.161	2.996	.003
	low social economic status	-2.196	1.291	-.095	-1.701	.090
	middle social economic class	-3.472	1.315	-.147	-2.641	.009
3	(Constant)	97.588	32.277		3.024	.003
	pulse frequency (beats/min)	.464	.781	.313	.594	.553
	low social economic status	-2.223	1.296	-.096	-1.715	.087
	middle social economic class	-3.432	1.324	-.146	-2.592	.010
	squared pulse rate	-.001	.005	-.153	-.290	.772

a Dependent Variable: systolic blood pressure (mmHg)

Model 3:

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.258(a)	.067	.058	16.86281
2	.284(b)	.081	.066	16.78795

a Predictors: (Constant), middle social economic class, pulse frequency (beats/min), low social economic status
b Predictors: (Constant), middle social economic class, pulse frequency (beats/min), low social economic status, mid_pulse, low_pulse

ANOVA(c)

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	6572.496	3	2190.832	7.705	.000(a)
	Residual	92130.781	324	284.354		
	Total	98703.277	327			
2	Regression	7952.374	5	1590.475	5.643	.000(b)
	Residual	90750.903	322	281.835		
	Total	98703.277	327			

a Predictors: (Constant), middle social economic class, pulse frequency (beats/min), low social economic status
b Predictors: (Constant), middle social economic class, pulse frequency (beats/min), low social economic status, mid_pulse, low_pulse
c Dependent Variable: systolic blood pressure (mmHg)

Coefficients(a)

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	106.752	6.625		16.113	.000
	pulse frequency (beats/min)	.239	.080	.161	2.996	.003
	low social economic status	-2.196	1.291	-.095	-1.701	.090
	middle social economic class	-3.472	1.315	-.147	-2.641	.009
2	(Constant)	101.067	7.155		14.125	.000
	pulse frequency (beats/min)	.306	.086	.207	3.568	.000
	low social economic status	12.841	9.568	.554	1.342	.181
	middle social economic class	6.729	9.051	.285	.743	.458
	low_pulse	-.181	.115	-.653	-1.578	.116
	mid_pulse	-.123	.109	-.435	-1.130	.259

a Dependent Variable: systolic blood pressure (mmHg)

In order to look at the crude differences in mean systolic blood pressure between the three groups, model 1 is fitted. Study the output of model 1, notice in particular how the independent variables are coded (***LOW:1 0 -1 and MID: 0 1 -1***), and answer questions (a) to (d).

a. What is the interpretation of the estimated regression coefficients of the independent variables "low social economic status" and "middle social economic status"? (1 mark)
   i. Give also the interpretation of the estimated intercept. (1 mark)
   ii. Compute the estimates for the mean systolic blood pressure of the three SES classes. (1 mark)
b. Are there significant differences in mean systolic blood pressures between the SES groups?
   i. Formulate the null hypothesis and give the p-value. (1 mark)
c. Give the estimate of the within groups standard deviation of systolic blood pressure. (1 mark)
   i. How can this be used to compute an (approximate) 95% confidence interval for the group means? (1 mark)
   ii. Give this confidence interval for the low SES group. (the number of individuals in the lower SES group was 138) (1 mark)
d. Give the estimate of the percentage variability in systolic blood pressure that is explained by differences between SES classes. (1 mark)

In order to look at the differences in mean systolic blood pressure between the SES groups corrected for pulse rate, model 2 was fitted. Study the output of model 2 and answer the questions (e) and (h).

e. Are there significant differences in mean systolic blood pressures between the SES groups corrected for pulse rate? (1 mark)
   i. Formulate the null hypothesis and give the p-value. (2 marks)
f. Give the estimate of the pulse rate corrected difference in mean systolic blood pressure between the low and middle SES group. (1 mark)
   i. Do the same for the low and high group and for the middle and high group. (1 mark)
g. Compute the estimate, based on model 2, of the mean systolic blood pressure for middle class people with pulse rate equal to 70 (1 mark)
h. One of the assumptions underlying model 2 is that the relation between systolic blood pressure and pulse rate is linear. Is this assumption reasonable in this case? (1 mark)
   i. Motivate your answer. (1 mark)

One of the assumptions of model 2 is that there is no interaction between SES classes and pulse rate. In order to investigate whether this assumption is justified, model 3 was fitted. Study the output of model 3 and answer the following questions.

i. Is there statistical evidence that there is interaction between SES class and pulse rate? (1 mark)

    i. Motivate your answer. (1 mark)

j. Give the equation of the estimated regression line (based on model 3) of systolic blood pressure against pulse rate for the low SES group. (1 mark)

What is the estimated difference (based on model 3) in mean systolic blood pressure (1 mark)

5. Question five (20 marks).

The table below gives results of 6 clinical trials comparing the risk of OHSS (ovarian hyperstimulation syndrome) between recombinant FSH and urinary FSH used during an IVF (in vitro fertilization) treatment.

Trial	No. of patients Rec FSH	No. of patients Ur FSH	OHSS Rec FSH	OHSS Ur FSH
1	585	396	19	8
2	57	33	3	0
3	54	35	2	1
4	119	114	6	2
5	60	63	2	1
6	105	67	8	3

A meta-analysis was carried out using Mantel-Haenszel's procedure, stratified on trial. Some SPSS output is given at the following pages. Read this output and answer the following questions.

Risk Estimate

Trial		Value	95% Confidence Interval	
			Lower	Upper
1	Odds Ratio for FSH (Recombinant / Urinary)	.614	.266	1.417

	For cohort OHSS = no	.987	.967	1.008
	For cohort OHSS = yes	1.608	.711	3.636
	N of Valid Cases	981		
2	For cohort OHSS = no	.947	.891	1.007
	N of Valid Cases	90		
3	Odds Ratio for FSH (Recombinant / Urinary)	.765	.067	8.765
	For cohort OHSS = no	.991	.918	1.071
	For cohort OHSS = yes	1.296	.122	13.763
	N of Valid Cases	89		
4	Odds Ratio for FSH (Recombinant / Urinary)	.336	.066	1.702
	For cohort OHSS = no	.967	.921	1.014
	For cohort OHSS = yes	2.874	.592	13.947
	N of Valid Cases	233		
5	Odds Ratio for FSH (Recombinant / Urinary)	.468	.041	5.297
	For cohort OHSS = no	.982	.928	1.039
	For cohort OHSS = yes	2.100	.195	22.561
	N of Valid Cases	123		
6	Odds Ratio for FSH (Recombinant / Urinary)	.568	.145	2.223
	For cohort OHSS = no	.967	.897	1.043
	For cohort OHSS = yes	1.702	.468	6.188
	N of Valid Cases	172		

Tests of Homogeneity of the Odds Ratio

	Chi-Squared	df	Asymp. Sig. (2-sided)
Breslow-Day	1.507	5	.912
Tarone's	1.507	5	.912

Mantel-Haenszel Common Odds Ratio Estimate

Estimate			.513
ln(Estimate)			-.668
Std. Error of ln(Estimate)			.308
Asymp. Sig. (2-sided)			???
Asymp. 95% Confidence Interval	Common Odds Ratio	Lower Bound	???
		Upper Bound	???
	ln(Common Odds Ratio)	Lower Bound	???
		Upper Bound	???

.

k. Make a 2X2 table for the first trial.

i. Compute the OHSS odds ratio of recombinant FSH treatment relative to urinary FSH treatment. (1 mark)
ii. Compute also the corresponding relative risk (1 mark)
iii. How are these estimates related to the estimates given for trial 1 in the first table of the SPSS output (1 mark)
iv. What is the difference between the two relative risk estimates? (1 mark)

l. Give the OHSS odds ratios of recombinant FSH relative to urinary FSH per trial. Is the assumption that the true odds ratios are equal across trials warranted? (7 marks)
   i. Motivate your answer. (1 mark)

m. Give the Mantel-Haenzel estimate of the common OHSS odds ratios of recombinant FSH relative to urinary FSH. (2 marks)
   i. Is it justified to interpret it as a relative risk? (1 mark)

n. Fill in the question marks in the third table. (5 marks)