COMPREHESIVE EXAM SAMPLE QUESTIONS

The following is a set of sample questions drawn from previous SPPH comprehensive exams. The sole intention of providing these sample questions is to give a sense of what the questions on the exam tend to look like. These sample questions DO NOT span the entire examinable content of the exam, and they do not reflect the length of the exam. They do not display every 'type' of question that may be asked. There may be redundancies in the content covered in these particular questions. In other words, they are simply a set of questions pulled from previous exams, and placed into this document, so that you have a sense of what sorts of questions are asked on the exam, and they sort of format they may take.

SAMPLE QUESTION:

Please describe how each of the following scenarios (a-d) might alter the observed odds ratio in case-control studies examining the relationship between a risk factor (low physical activity) and a disease (heart disease). For each answer, please explain (qualitatively, no numbers) which cell(s) in a hypothetical 2x2 table would change and in which direction (Maximum 2 sentences per scenario)

a) Controls selected from a population that tends to be much healthier in terms of exposure to risk factors (lower level of exposure) than the population from which the cases are drawn.

b) Poor recall of exposure, non-differential with respect to disease status (i.e., equally poor recall of exposure among cases and controls).

c) Early signs of disease that lead to reductions in exposure (i.e., reductions in 'highrisk' behaviours) immediately prior to diagnosis of disease.

d) Cases recall more exposure than actually occurred.

SAMPLE QUESTION:

In survival analysis, one important assumption is that censoring is non-informative. In other words, we assume that not having observed the event is not related to the probability of the event occurring. Suppose that censoring was related to the probability of the event occurring. In other words, censoring was informative. Create an example to demonstrate how this may bias survival upward, or downward. The example does not have to be numeric, but it should refer to a specific disease and outcome, and explain the reason why censoring is informative.

SAMPLE QUESTION:

Below is output from an Analysis of Covariance (ANCOVA) and a Linear Regression model, with categorical and continuous independent variables on a simple random sample of cafeteria employees.

This is a statistical model examining the relationships between retirement savings and future retiree characteristics. In the model, the dependent variable is retirement account savings (measured in thousands). The independent variables are: high school completing (Y/N) and age (numeric).

The ANCOVA table is:

Source	DF	Sum of Squares	Mean Square	F-stat	P-value
Model	2	100	50	13	0.001
Error	23	92	4		
Total	25	192			

The Model Parameters are:

Coefficient	Estimate	Standard Error	t-stat	P-value
Intercept	17	5	3.4	0.001
High School (N)	-4.0	1	4	0.001
High School (Y)	0.0	.	.	.
Age	1.5	0.5	3	0.030

Please use the above to answer the following questions:

1. Please respond to the following:
(a) Does completing high school have an effect on retirement savings?
(b) Describe the relationship between high school and pension savings

2. What is the expected retirement savings of a 30 year old who did not complete high school?

3. What is the expected difference in savings for two individuals that are 10 years apart in age?

4. For someone 40 years old and finished high school, what is the approximate probability that pension savings exceeds 81 thousand?

5. What is the approximate probability that the difference in retirement savings between person A and person B is greater than 5 thousand, when person A did not finish high school and is 50 years old, while person B did finish high school and is 40 years old

SAMPLE QUESTION:

A researcher is investigating the hypothesis that liver cirrhosis is associated with heavy alcohol drinking using a case-control design. The table below shows levels of alcohol use for cases (individuals with liver cirrhosis) and controls. Cases were obtained from a disease registry. The controls were obtained from a hospital emergency department admissions registry. Cases and controls completed a questionnaire and reported habitual alcohol use.

Self-reported alcohol use	Cases	Controls
Heavy	165	85
Light	35	115

(a) [1 mark] Calculate the OR for heavy alcohol use in cases compared to controls.
   (expected answer: calculation)

(b) [3 marks] Discuss briefly the main type of bias that may affect this result.
   (expected answer: 2-3 sentences)

(c) [3 marks] Describe the likely impact of this type of bias on the OR.
   (expected answer: 2-3 sentences)

(d) [3 marks] Describe your approach to mitigating this bias.
   (expected answer: a few short sentences)

SAMPLE QUESTION

Data from the 1996 General Social Survey was used to study the relationship between the number of volunteer activities participated in per year, and various other variables. The variables recorded were:

Y = Number of volunteer activities in 1 year
$X1$ = Gender (1=female, 0=male)
$X2$ = VisMinority (1=visible minority, 0=not visible minority)
$X3$ = Education (categorized as A=below high school, B=high school,
       C=undergraduate, D=post graduate)
$X4$ = Income (measured in units of $10,000)

The researchers used a Poisson regression model for number of volunteer activities in a year, including a quadratic term for income to address non-linearity. The model coefficients are below (expressed in terms of the log-number of volunteer activities).

	Coefficient	SE of Coefficient
Constant/Intercept	-1.4	0.60
Gender(female)	+0.93	0.22
VisMinority(yes)	-0.67	0.17
Education(B)	+0.33	0.13
Education(C)	+1.31	0.16
Education(D)	+1.49	0.21
Income	-0.117	0.03

Income^2	+0.011	0.002
Gender(female)*VisMonority(yes)	+1.83	0.24

(a) [2 marks] Based on the fitted model, state the effect each variable has on the expected number of volunteer activities (ie) does it statistically significantly increase/decrease the expected outcome?
   (expected answer: a few short sentences)

(b) [5 marks] Based on the model, what is the expected number of volunteer activities in a year for a female, who is a visible minority in education category C, with an income of $35,000?
   (expected answer: numeric answer, showing work)

(c) [6 marks] What is the probability that an individual as described in part (b) (female, visible minority in education category C with income of $35,000) participates in at least 3 volunteer activities in a year?
   (expected answer: numeric answer, showing work)

(d) [8 marks] Report and interpret relevant rate ratio(s) (RR) for comparing the number of volunteer activities for males and females.
   (expected answer: numeric answers, showing work. Few short sentences)

(e) [4 marks] Report and interpret relevant rate ratio(s) (RR) for comparing the number of volunteer activities for someone in education category D to education category C. Do you believe this RR to be statistically significant? Make sure to justify your answer.
   (expected answer: numeric answers, showing work. Few short sentences)

SAMPLE QUESTION

Provide an interpretation of the data presented in the following table:
   (a) From a population and public health perspective
   (b) From an individual level perspective such as in genetic counseling
(expected answer: half a page total)

Incidence of Down's syndrome according to maternal age.[7]

Maternal age (years)	Risk of Down's syndrome per 1000 births	Total births in age group (as % of all ages)	% of total Down's syndrome occurring in age group
< 30	0.7	78	51
30–34	1.3	16	20
35–39	3.7	5	16
40–44	13.1	0.95	11
> 45	34.6	0.05	2
All ages	1.5	100	100