



KENYATTA UNIVERSITY

UNIVERSITY EXAMINATIONS 2010/2011

INSTITUTION BASED PROGRAMME (IBP) AUGUST SESSION

EXAMINATION FOR THE DEGREE OF MASTER OF SCIENCE

SMA/SS/569: OBSERVATION STUDIES

DATE: THURSDAY, 29TH DECEMBER 2011

TIME: 9.00 A.M. – 12.00 P.M.

INSTRUCTIONS: Answer any THREE Questions.

1. (a) What is observational studies. Compare observational studies with controlled experimentation.
- (b) Define the term “confounding variable” using an examples explain this term.
- (c) The following data gives the fatalities in matatus.

Table 1.

Number of	Treatment Seatbelt	
	Worn	Not worn
Drivers killed	10	20
Drivers not killed	40	30

Table 2.

Number of	Low Impact		High Impact	
	Seat belt		Seat belt	
	Worn	Not worn	Worn	Not known
Drivers killed	4	2	6	18
Drivers not killed	36	18	4	12

Analyse the two set of data. Can you conclude that seatbelts saves life.

- (d) Let Y denotes the blood alcohol level of a driver after television advertisement campaign (Y_{1i}) and before the campaign (Y_{0i}) for different ages.

Age	% ($Y_{1i} > 80$)	% ($Y_{0i} > 80$)
17-25	20	15
25-35	15	15
35-45	10	10
45	10	15

Using the three measures of risk determine whether the T.V campaign was a success.

2. (a) Define the term “propensity score” $e(x)$ for a vector covariate $\tilde{x}$ and a binary treatment.

- (b) If treatment is ignorable given $\tilde{x}$ show that $(Y_0, Y_1) \perp\!\!\!\perp Z \mid \tilde{x} \Rightarrow (Y_0, Y_1) \perp\!\!\!\perp e(\tilde{x})$.

- (c) Explain what is Lord’s Paradox.

Let $y(t_1)$ and $y(t_2)$ be the weights for males and females at time t_1 and t_2 .

Give the estimates of the treatment effect using

- (i) Gain score
(ii) Analysis of Covariance.

When are the two treatment effects equal.

3. (a) Define the term “Matching Methods” and standardization methods”. What is the main difference between them.
(b) Stating the underlying assumptions, describe the mean and pair matching methods.
(c) If the regressions are quadratic that is

$$E(Y_1) = \alpha_1 + x\beta + x^2r$$

$$E(Y_0) = \alpha_0 + x\beta + x^2r$$

Show that in case of

- (i) mean matching

$$E(\bar{y}_T - \bar{y}_c \mid x) = \alpha_1 - \alpha_0 + (S_T^2 - S_c^2)r$$

$$\text{where } S_c^2 = E(X_c^2) - [E(\bar{x}_c)]^2$$

$$S_T^2 = E(x_T^2) - [E(\bar{x}_T)]^2$$

- (ii) Pair matching

$$E(\bar{y}_T - \bar{y}_c | x) = \alpha_1 - \alpha_0.$$

Which of the two matching methods is preferable and why?

4. (a) Define and differentiate between Analysis of Variance (ANOVA) and Analysis of Covariance (ANCOVA).

- (b) What is a confounding variable?

Suppose $\tilde{x}$ is a continuous confounding variable where

$$Y_1 = \alpha_1 + X\beta + \epsilon$$

$$Y_0 = \alpha_0 + X\beta + \epsilon$$

- (i) For given $X = x$ show that

$$E(\bar{Y}_T - \bar{Y}_c | x) = \alpha_1 - \alpha_0$$

i.e. Treatment effect is unbiased.

- (ii) Assume that treatments is ignorable given X

$$\text{i.e. } (Y_0, Y_1) \perp\!\!\!\perp Z | x$$

then show that

$$E(\bar{Y}_T - \bar{Y}_c | X) = \alpha_1 - \alpha_0 + (\bar{X}_T - \bar{X}_c)\beta$$

i.e. Treatment is biased.

- (c) The following results are from a study by Field (2009) on Viagra. The partner's libido is measured in how often they tried to initiate sexual contact. This partners libido is taken as the covariate variable.

Tables 1, 2 and 3 are results from SPSS output.

Table 1: (mean and standard deviation)

Dose	Participant's Libido	Partner's Libido
Placebo	3.22 (1.79)	3.44 (2.07)
Low Dose	4.88 (1.46)	3.12 (1.73)
High Dose	4.85 (2.12)	2.00 (1.63)

Standard deviation in parenthesis.

Table 2: ANOVA (Partner Libido excluded)

Dependent Variable Libido

Source of variation	d.f	Sum of squares	Mean sum of squares	F-ratio	Sig P-value
intercept	1	535.184	535.184		
Dose	2	16.84	8.422	2.41	0.108
Error	27	94.123	3.486		
Total	30	683			

Table 3: ANCOVA (Partners Libido included)

Dependent variable: Libido

Source	D.F	S. Square	N.S.S.	F	P-value
Intercept	1	76.069	76.069	25.02	0.000
Partners Libido	1	15.076	15.076	4.959	0.035
Dose	2	25.185	12.593	4.142	0.027
Error	26	79.047			
Total	30	683.000			

Is there a significant effect on Viagra on levels of libido after controlling for the effect of the covariate (partners libido).

5. (a) Define the Hawthorne reactive effect. Give four situation where this effect is present.
- (b) Differentiate between internal and external validity. Explain how Hawthorne reactive effect affect each.
- (c) (i) Briefly describe one method of adjustment for the selection bias.
(ii) Data on death rates for two populations
Population A: Control group
Population B: Treatment group

Age	Population A		Population B	
	Number (n_{cj})	Death rate $\bar{y}_{cj}$	Number (n_{Tj})	Death rate ($\bar{y}_{Tj}$)
15-34	76	1	90	3
35-44	8	25	62	18
45-54	5	75	47	43
55-64	3	120	31	71
65-74	2	140	18	89
75+	0	303	8	138

Compute the standardization rates for the above data confounding on the age and interpret the results.
