Exercise for Course 607

based on "Acute mountain sickness in western tourists around the Thorong Pass (5400 m) in Nepal" by B.Kayser in Journal of Wilderness Medicine 2, 110-117 (1991)

INTRODUCTION

1) What was the aim of this study? What is the outcome variable? Is it a continuous or a binary outcome?

Answer:
The aim of this study was to describe the prevalence of acute mountain sickness over a particular mountain pass in Nepal. The outcome variable was the proportion (%) of trekkers having AMS, which means having a predefined score or higher on one or two parts of the Environmental Statistical Questionaire (III). This a binary outcome variable (AMS yes or no).

METHODS

2) The author describes the scoring system of the AMS-C part and the AMS-R part.

a) What is the range of the scores, what is defined as "threshold"?
b) Sketch a probable frequency distribution for the AMS-C scores.
c) What will an AMS-C score of 4.6 probably indicate? Is the difference between an AMS-C score of 0.9 and a score of 4.6 of interest for this study?

Answer:
a) Range for AMS-C score and AMS-R: 0-5. Threshold for AMS-C .7, for AMS-R .6
b) Frequency distribution: probably median around .5 , skewed to the right.
c) AMS-C score of 4.6 will probably indicate severe AMS (high altitude cerebral edema). No, the difference between an AMS-C score of 0.9 and 4.6 is not of interest in this study, since the threshold for AMS-C was 0.7

RESULTS

3) In table 2 the author reports in the first column the mean scores for AMS-C for women and men.

a) Comparison between the two scores is made with a t-test. Comment briefly on the use of the t-test in this case. Suggest another testing method.
b) State the null and alternative hypotheses and show that the exact t-value is 2.90 using pooled sample variances (the author probably did not do so). Explain what means p-value<0.01.
c) You have been taught in your course that not p-values but confidence intervals should be reported. Please calculate a 95% confidence interval for the author(use same pooled variances) and explain him what this means.

Answer

a) If the author uses the t-test for comparisons of means we have to assume symmetric distributions for AMS-C scores in his sample (it is not reported and doubtful). Alternative (if no symmetric distribution): Wilcoxon's rank sum test.
`Comment by jh... but large n's -> CLT -> ok with t`

b) H0: $\mu_{women} = \mu_{men}$ (or H0: $\mu_{women} - \mu_{men} = 0$) versus Ha: $\mu_{women} \neq \mu_{men}$. Pooled sample variance is 0.55. t test for difference in means = $(0.68-0.91) / \sqrt{0.55/193 + 0.55/160} = -2.90$ with df=193+160–2 = 351 p-value (1-sided): $.001<p<.0025$, p-value 2 sided $.002<p<.005$ so $p<.01$

p-value <.01 means that if there is no difference in AMS-C score between men and women ,the probability of finding this difference or more extreme is smaller than 1%.

c) t , df=100 from table, "safe"=1.984 : 95% CI for difference in mean AMS-C score for men and women: (.91-.68)±1.984* square root of [(0.55/193) + (0.55/160)= [.073; .387]
Meaning: we can be 95% confident that the true population parameter for difference in mean AMS-C score will be between .073 and .387.

4) In the second column of table 2 data are given on AMS-C prevalence. The difference between men and women is said to be not-significant.
a) Is this in in contradiction with the result you got in question 3?
*b) Derive the 2*2 table for prevalence of AMS-C in men and women and calculate the Chi square. Show that you get the same p-value using a z-test (use more exact proportions for men and women to avoid rounding differences). Do you think that if continuity correction was used, this difference could have reached significance?*
c) What is the RR for difference in prevalence for AMS-C in men and women, and its corresponding 95% CI?

Answer
a) No, there is no contradiction between a significant difference in (absolute) AMS-C score and a non significant difference in AMS-C prevalence between men and women. (Even if there would have been no AMS at all in both groups the difference in absolute score could have been significant.)
b)

	men	women	
AMS-C +	75 (38.86%)	78 (48.75%)	153
AMS-C -	118	82	200
	193	160	353

Chi-square (by formula from book)=3.48 p-value (table): (df=1) $.05<p<.10$ (epi-info: p=0.062)

Exercise for Course 607

based on "Acute mountain sickness in western tourists around the Thorong Pass (5400 m) in Nepal" by B.Kayser in Journal of Wilderness Medicine 2, 110-117 (1991)

for z -score: H0: (πmen=π(women versus Ha: (πmen≠πwomen)   p (overall)= 153/353=43.3%.

z= ((.3886.-.4875)-0) / √ [.433*.567*(1/193+1/160)]=-1.867 p-value (2-sided): 2*0.0309=0.0618

If continuity correction had been performed p-value would be slighly higher, so definitely not significant (continuity correction leads to more "conservative" p-values.)

c) RR= (75/193) / (78/160)=0.797 95% CI (test-based method): 0.797 (1±1.96√3.48)= [.628; 1.012]

5) One of the referees is interested in the relation between BMI and AMS-C score in men. He asks the author to analyse those results by simple lineair regression.

a) Please do this for the author. Sketch the line, assuming the average BMI in men is 22.7 kg/m2 , the average AMS-C score is 0.68 and the slope (regression coefficient) for BMI is 0.05.

b) Explain what means "slope for BMI is 0.05". In what units is this regression coefficient expressed, and can its significance be tested? If yes, explain the principle, and what will be the null hypothesis? If no, why not?

Answer:

a) Graph with BMI (independent variate) on X-ax, AMS-C score (dependent variate) on Y-ax, with slope 0.05, going through point (x bar, y bar)=(22.7, 0.68).

b) Regression coefficient for BMI is 0.05 means that with every unit increase in BMI, the average AMS-C score will increase by 0.05. Its units=units of AMS-C/unit of BMI=1/ (kg/m2)=m2/kg
Its significance can be tested by a t-test, null hypothesis is (=0 versus Ha: ((0.t=(0.03-0)/SE(^(), df=193-2=191

6) A new finding in this study was the difference in AMS prevalence between men and women. The author wants to repeat his study to re-investigate this finding. He asks your advice about the sample size needed. He considers a difference of 10% in prevalence rates of "clinical significance". Calculate the sample size, assuming equal n in both groups, (=0.05, power of 80%, using overall AMS prevalence rate of 63%.
If he redid his study on another (lower) mountain pass where the overall AMS prevalence rate was only 50%, would he need more or fewer subjects?

Answer:

n per group=16 x 0.63 x 0.37 / 0.10^2 =372.96 so 373 per group If overall prevalence of AMS 50% he would need more subjects (400).

```
Comment by jh... wouldn't necessarily use 63% from
this study... might use 50% to be on safe side. Also,
63% is subject to sampling variation.
```

exercise prepared by Barbara Broërs , June 1995