Advanced Qualitative Data Analysis
STATISTICS 455
Noelle I. Samia, Ph.D.

Homework # 2:

- Suppose that $Y = (Y_1, Y_2, Y_3)$ is multinomial $(n, p = (p_1, p_2, p_3))$.
  You are to test $H_0 : p_1 = p_2 = p_3$ using S_R^2 or S_U^2, where
  S_R^2 is the score test statistic for H_0 versus the restricted alternative $H_1 : p_1 = p_2 \neq p_3$;
  and S_U^2 is the score test statistic for H_0 versus the unrestricted alternative $H_1^* :$ not H_0.
  Recall that $S_R^2 \sim A\mathcal{X}^2(1, \lambda_R)$
  and $S_U^2 \sim A\mathcal{X}^2(2, \lambda_U)$.
  Let $P_R = P(S_R^2 \geq 3.8415),\ aP_R = P(\mathcal{X}^2(1, \lambda_R) \geq 3.8415)$;
  and $P_U = P(S_U^2 \geq 5.9915),\ aP_U = P(\mathcal{X}^2(1, \lambda_U) \geq 5.9915)$;
  where 3.8415 is the 95^{th} quantile of the $\mathcal{X}^2(1)$ distribution
  and 5.9915 is the 95^{th} quantile of the $\mathcal{X}^2(2)$ distribution.
  The values aP_R and aP_U are the chi-squared approximations to the exact tail probabilities P_R and P_U, respectively.

  a. Write explicitly S_R^2 and S_U^2 in terms of Y_1, Y_2, and n.

  b. Complete the following table, and comment on your findings.
  Hint: Use Monte Carlo simulation to compute exact probabilities P_R and P_U.

  Note that the table, once it is filled in, gives the Monte Carlo estimates and chi-squared approximations for the probabilities of rejecting H_0 for different values of p_T, using an approximate 0.05 size test; i.e. these are 'power' values.

Sample size n	True probability p_T	P_R	aP_R	P_U	aP_U
75	$(1/3, 1/3, 1/3)$				
75	$(1/4, 1/4, 2/4)$				
75	$(1/6, 3/6, 2/6)$				
75	$(0.2, 0.3, 0.5)$				
250	$(1/3, 1/3, 1/3)$				
250	$(0.3, 0.3, 0.4)$				
250	$(0.22, 0.4467, 0.3333)$				
250	$(0.250, 0.300, 0.450)$				
250	$(0.22, 0.40, 0.38)$				

  c. Let n_R be the sample size needed so that the approximate power based on S_R^2 is 0.8; i.e. $P(\mathcal{X}^2(1, \lambda_R) \geq 3.8415) = 0.8$.
  Let n_U be the sample size needed so that the approximate power based on S_U^2 is 0.8; i.e. $P(\mathcal{X}^2(2, \lambda_U) \geq 5.9915) = 0.8$. Fill in the following table of required sample sizes, and comment on your findings.

True probability p_T	n_R	n_U
$(1/3, 1/3, 1/3)$		
$(1/4, 1/4, 2/4)$		
$(1/6, 3/6, 2/6)$		
$(0.2, 0.3, 0.5)$		
$(0.3, 0.3, 0.4)$		
$(0.22, 0.4467, 0.3333)$		
$(0.250, 0.300, 0.450)$		
$(0.22, 0.40, 0.38)$		