HW2

Brody Kendall

10/14/2021

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

Note that $\hat{\pi}_{0,i} = (E[Y_i] \text{ under } H_0) = 1/3$ for $i = 1, 2, 3$. Also, $Y_3 = n - Y_1 - Y_2$
Finding S_U^2 is relatively straightforward:
$S_U^2 = \Sigma_{i=1}^3 \frac{(Y_i - \hat{m}_{0,i})^2}{\hat{m}_{0,i}}$ where $\hat{m}_{0,i} = n\hat{\pi}_{0,i} = n/3$ for $i = 1, 2, 3$
This gives $S_U^2 = \frac{(Y_1-n/3)^2+(Y_2-n/3)^2+(n-Y_1-Y_2-n/3)^2}{n/3}$
$= \frac{6Y_1^2}{n} + \frac{6Y_2^2}{n} + \frac{6Y_1Y_2}{n} - 6Y_1 - 6Y_2 + 2n$

Now to find S_R^2:
$\Omega_1 = \{p \in \Omega : p_1 = p_2\} = \{p : p = (\beta, \beta, 1 - 2\beta), \beta \in \mathcal{B}\}$ for $\mathcal{B} = [0, 1/2]$
$dim(\Omega_1) = 1$
$\Omega_0 = \{p = (1/3, 1/3, 1/3)\}$
$dim(\Omega_0) = 0$
$\nu = 1 - 0 = 1$
$L(\beta|y) = (y_1 + y_2)log\beta + y_3log(1 - 2\beta)$
$s(\beta|y) = \frac{y_1+y_2}{\beta} - \frac{2y_3}{1-2\beta}$
$\frac{\partial}{\partial\beta}s(\beta|y) = -\frac{y_1+y_2}{\beta^2} - \frac{4y_3}{(1-2\beta)^2}$
$B(\beta) = \frac{2n\beta}{\beta^2} + \frac{4n(1-2\beta)}{(1-2\beta)^2} = \frac{2n}{\beta(1-2\beta)}$
Now, $S_R^2 = S^2(\Omega_0|\Omega_1) = \frac{s(\hat{\beta}_0|Y)^2}{B(\hat{\beta}_0)}$ where $\hat{\beta}_0 = (1/3, 1/3, 1/3)$
$S_R^2 = \frac{81(Y_1+Y_2-2n/3)^2}{18n} = \frac{9(Y_1+Y_2-2n/3)^2}{2n}$

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

$\tilde{y} = E_{p_t}(Y) = (np_1, np_2, np_3)$
$\lambda_R = S_R^2(\tilde{y}) = \frac{9(np_1+np_2-2n/3)^2}{2n}$
$\lambda_U = S_U^2(\tilde{y}) = \frac{6(np_1)^2}{n} + \frac{6(np_2)^2}{n} + \frac{6(np_1)(np_2)}{n} - 6(np_1) - 6(np_2) + 2n$
For the first row, we get $\tilde{y} = (1/3, 1/3, 1/3)$ so
$\lambda_R = \frac{9(75(1/3)+75(1/3)-2(75)/3)^2}{2(75)} = 0$ and

$\lambda_U = \frac{6(75*1/3)^2}{75} + \frac{6(75*1/3)^2}{75} + \frac{6(75*1/3)(75*1/3)}{75} - 6(75*1/3) - 6(75*1/3) + 2(75) = 0$, as expected.

Now, we can use these λ's to calculate aP_R and aP_U in R using the pchisq() function. For the first row, we get $aP_R = P(\chi^2(1, \lambda_R = 0) \geq 3.8415) \approx 0.05$ and $aP_U = P(\chi^2(2, \lambda_U = 0) \geq 5.9915) \approx 0.05$

To estimate P_R and P_U we will use Monte Carlo simulation. For each row, we will first generate 10,000 random samples from the given $multinomal(n, p_1, p_2, p_3)$ distribution. Then we will calculate S_R^2 and S_U^2 for each of these 10,000 samples. Our estimate for P_R will be the proportion of samples that result in a S_R^2 value greater than 3.8415. Similarly, our estimate for P_U will be the proportion of samples that result in a S_U^2 value greater than 5.9915.

Sample size n	True probability p_T	P_R	aP_R	P_U	aP_U
75	(1/3, 1/3, 1/3)	0.04	0.05	0.05	0.05
75	(1/4, 1/4, 2/4)	0.83	0.86	0.79	0.79
75	(1/6, 3/6, 2/6)	0.03	0.05	0.92	0.90
75	(0.2, 0.3, 0.5)	0.82	0.86	0.84	0.83
250	(1/3, 1/3, 1/3)	0.05	0.05	0.05	0.05
250	(0.3, 0.3, 0.4)	0.64	0.61	0.50	0.50
250	(0.22, 0.4467, 0.3333)	0.05	0.05	0.99	0.98
250	(0.250, 0.300, 0.450)	0.97	0.97	0.95	0.96
250	(0.22, 0.40, 0.38)	0.37	0.35	0.96	0.94

It is notable how much small changes in the true probability alter the P's and aP's, especially when we consider the vast differences between these values for the restricted and unrestricted tests. For example, the two rows where the true probability is $(1/3, 1/3, 1/3)$, all the values are 0.05 (or very close to 0.05). However, when the true probability is $(1/6, 3/6, 2/6)$, we get similar values for P_R and aP_R (close to 0.05) but P_U and aP_U are much higher - both at least 0.9. This shows how much accurate knowledge about the field in question can help when constructing tests like the ones above. At the same time, though, it shows how much inaccurate knowledge can hinder these types of tests.

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. $P(X^2(2, \lambda_U) \geq 5.9915) = 0.8$. Fill in the following table of required sample sizes, and comment on your findings.

To find n_R and n_U, we will use a systematic approach, instantiating both at 1 and increasing by 1 at each iteration until the desired power is achieved. We will cap both n_R and n_U at 10,000 and assume that if this cap is achieved, there is no such value that results in the desired power, marked in the table as NA.

For example, the first row yields a non-centrality parameter of 0 for both the restricted and unrestricted tests regardless of n. Therefore, $P(\mathcal{X}^2(1, \lambda_R = 0) \geq 3.8415) = 0.05 \neq 0.8$ and $P(X^2(2, \lambda_U = 0) \geq 5.9915) = 0.05 \neq 0.8$, regardless of n_R and n_U.

True probability p_T	n_R	n_U
(1/3, 1/3, 1/3)	NA	NA
(1/4, 1/4, 2/4)	63	78
(1/6, 3/6, 2/6)	NA	58
(0.2, 0.3, 0.5)	63	69
(0.3, 0.3, 0.4)	393	482
(0.22, 0.4467, 0.3333)	NA	125
(0.250, 0.300, 0.450)	129	149
(0.22, 0.40, 0.38)	801	165

Again, we see a great difference between the results for the restricted tests vs the unrestricted tests depending on the true probability. Note that when the alternate restricted hypothesis is true, n_R is always smaller than n_U as expected. In the table, this is when $p_T = (1/4, 1/4, 2/4)$ or $(0.3, 0.3, 0.4)$. In addition, even when the alternate restricted hypothesis is close to true (as in p_2 is closer to p_1 than it is to p_3), n_R is still smaller than n_U. In the table, this is when $p_T = (0.2, 0.3, 0.5)$ or $(0.250, 0.300, 0.450)$. As stated before, when $p_T = (1/3, 1/3, 1/3)$, the non-centrality parameters for both the restricted and unrestricted tests are 0, so neither test can achieve the desired power. However, there are further probabilities ($p_T = (1/6, 3/6, 2/6)$ and $(0.22, 0.4467, 0.3333)$) where the restricted non-centrality parameter is 0, but the unrestricted non-centrality parameter is nonzero. This is because the unrestricted alternate hypothesis is true, but p_3 is equidistant from p_1 and p_2. The final row is now the only one which we have not considered. In this case, p_2 and p_3 are very close but both are relatively far from p_1, so the restricted alternate hypothesis is very close to being false while the unrestricted alternate hypothesis is certainly true. This results in a much lower n_U than n_R. As in part (b), this table shows how useful accurate knowledge can be when constructing tests as well as the detrimental effects of inaccurate knowledge.

I have submitted this pdf along with the R file containing the functions I used to calculate the various powers and n's.