Statistics 581, Problem Set 3 Solution supplement

Wellner; 10/21/2005

1. Ferguson, ACILST, problem 1, page 65, modified.

   In a multinomial experiment with sample size $n = 100$ and 3 cells with null hypothesis $H_0 : \underline{p}_0 = (1/3, 1/3, 1/3)$, what is the approximate power at the alternative $\underline{p} = (.25, .5, .25)$ when the level of significance is $\alpha = .05$? $\alpha = .01$? How large a sample size is need to achieve power 0.8 at this alternative when $\alpha = .05$? $\alpha = .01$?

   (a) In a multinomial experiment with sample size 100 and 3 cells with null hypothesis $H_0 : p_1 = 1/4, p_2 = 1/2, p_3 = 1/4$, what is the approximate power at the alternative $p_1 = 0.2, p_2 = 0.6, p_3 = 0.2$ when the level of significance is $\alpha = 0.05$? $\alpha = 0.01$? (b) How large a sample size is needed to achieve power 0.9 at this alternative when $\alpha = 0.05$? $\alpha = 0.01$?

   Solution: Now

$$n^{1/2}(\underline{p}-\underline{p}_0) = 10((1/4, 1/2, 1/4)-(1/3, 1/3, 1/3)) = 10(-1/12, 1/6, -1/12) = (-5/6, 5/3, -5/6)\,,$$

so the non-centrality parameter is

$$\delta = \frac{(5/6)^2}{1/3} + \frac{(5/3)^2}{1/3} + \frac{(5/6)^2}{1/3} = 3 \cdot 25 \left\{ \frac{1}{36} + \frac{1}{9} + \frac{1}{36} \right\} = \frac{25}{2} = 12.5\,.$$

Thus the approximate power via $\chi_2^2(\delta)$ is

$$P(\chi_2^2(12.5) \geq \chi_{2,.05}) = P(\chi_2^2(12.5) \geq 5.991) = .896\,, \qquad \text{when} \quad \alpha = .05\,,$$

and

$$P(\chi_2^2(12.5) \geq \chi_{2,.01}) = P(\chi_2^2(12.5) \geq 9.210) = .744 \qquad \text{when} \quad \alpha = .01\,,$$

(b) Now we want to find n so that

$$P(\chi_2^2(\delta_n) \geq 5.991) = .90$$

where

$$\delta_n = n \left\{ \frac{(1/12)^2}{1/3} + \frac{(1/6)^2}{1/3} + \frac{(1/12)62}{1/3} \right) = n/8\,.$$

In this case we find that $\delta_n = n/8 = 12.6539$, so that $n = 8 \cdot 12.6539 \approx 102$. When $\alpha = .01$ we find that $\delta_n = n/8 = 17.4267$ so that $n = 8 \cdot 17.4267/(.04) \approx 140$.

The alternative approximation to power that we derived in class is

$$\begin{aligned} P_p(Q_n \geq \chi^2_{k-1,\alpha}) &= P_p\left(\sqrt{n}(n^{-1}Q_n - q) \geq \sqrt{n}(n^{-1}\chi^2_{k-1,\alpha} - q)\right) \\ &\doteq P\left(N(0, d^T A d) \geq \sqrt{n}(n^{-1}\chi^2_{k-1,\alpha} - q)\right) \\ &= 1 - \Phi(\sqrt{n}(n^{-1}\chi^2_{k-1,\alpha} - q)/\sqrt{d^T A d}) \end{aligned}$$

where $d \equiv 2\text{diag}(1/p_0)(p - p_0)$, $A = \text{diag}(p) - pp^T$, and $q = \sum_{j=1}^{k}(p_j - p_{j0})^2/p_{j0}$. In the present case I calculate $q = 1/8$, $d = (-1/2, 1, -1/2)^T = (-1, 2, -1)/2$, and

$$A = \text{diag}(p) - pp^T = \frac{1}{16}\begin{pmatrix} 3 & -2 & -1 \\ -2 & 4 & -2 \\ -1 & -2 & 3 \end{pmatrix}$$

so that $d^T A d = (3/4)^2$. Thus the approximation becomes

$$P_p(Q_n \geq \chi^2_{2,\alpha}) \doteq 1 - \Phi(\sqrt{n}(n^{-1}\chi^2_{2,\alpha} - 1/8)/(3/4)).$$

When I calculate I get

$$P_p(Q_n \geq \chi^2_{2,.05}) \doteq 1 - \Phi(\sqrt{n}(n^{-1}\chi^2_{2,.05} - 1/8)/(3/4)) = 0.807249$$
$$P_p(Q_n \geq \chi^2_{2,.01}) \doteq 1 - \Phi(\sqrt{n}(n^{-1}\chi^2_{2,.01} - 1/8)/(3/4)) = 0.669532,$$

which are both somewhat lower than suggested by the non-central chi-square approximation.

From a Monte-Carlo of the chi-square statistic under this particular alternative with $nreps = 5000$, I get the following empirical (or estimated powers)

$$\widehat{P_p}(Q_n \geq \chi^2_{2,.05}) = 0.893,$$
$$\widehat{P_p}(Q_n \geq \chi^2_{2,.01}) = 0.7298.$$

The following plot shows the empirical distribution function of the 5000 values of Q_n in this case with $n = 100$ and true $p = (1/4, 1/2, 1/4)$ together with the non-central chi-square approximation as derived in class and computed explicitly above in this case. It appears that the non-central chi-square approximation is quite accurate (and beats the normal theory approximation derived under a fixed alternative).

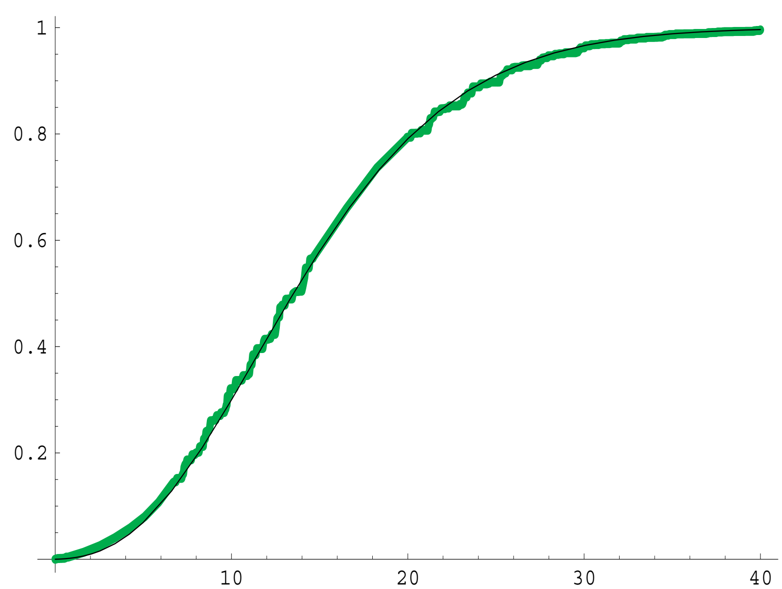

Figure 1: Plot of empirical of Monte-carloed Q_n values and non-central chi-square approximation