Problem 7 (AMS 205B) Solution:

(a) The joint density for $X_1, ..., X_n$ and $Y_1, ..., Y_m$ is given by

$$f(Y_1, ..., Y_m, X_1, ..., X_n | p_1, p_2) = p_1^{\sum_{i=1}^n X_i}(1-p_1)^{n-\sum_{i=1}^n X_i} p_2^{\sum_{j=1}^m Y_j}(1-p_2)^{m-\sum_{j=1}^m Y_j}.$$

Therefore, for any other realization $\mathbf{U} = (U_1, \ldots, U_n)$ of $\mathbf{X}$ and $\mathbf{V} = (V_1, \ldots, V_m)$ of $\mathbf{Y}$,

$$\frac{f(\mathbf{X}, \mathbf{Y}|p_1, p_2)}{f(\mathbf{U}, \mathbf{V}|p_1, p_2)} = \left(\frac{p_1}{1-p_1}\right)^{\sum_{i=1}^n X_i - \sum_{i=1}^n U_i} \left(\frac{p_2}{1-p_2}\right)^{\sum_{j=1}^m Y_j - \sum_{j=1}^m V_j}$$

This ratio is a constant function of the parameter vector if and only if ($\sum X_i = \sum U_i$, $\sum Y_j = \sum V_j$). Hence a minimal sufficient statistic is $(\sum X_i, \sum Y_j)$. Since, this is a two parameter exponential family and the parameter space contains an open rectangle in $(0,1) \times (0,1)$, the minimal sufficient statistic is also complete.

Note that, under the assumption that $p_1 = p_2$, $E(m \sum X_i - n \sum Y_j) = 0$, i.e. there exists a nonzero function of $(\sum X_i, \sum Y_j)$ whose expectation is zero for all parameters. Therefore, $(\sum X_i, \sum Y_j)$ is not complete sufficient under the additional assumption.

(b) Define,

$$H(\mathbf{X}) = \begin{cases} 1 \text{ if } \sum_{i=1}^{n-2} X_i > X_{n-1} + X_n \\ 0 \text{ o.w.} \end{cases}$$

Clearly, $E[H(\mathbf{X})] = h(p_1)$, i.e. $H(\mathbf{X})$ is an unbiased estimator of $h(p_1)$. We will use Rao-Blackwell theorem to find a better estimator than $H(\mathbf{X})$ as following.

$$E[H(\mathbf{X})| \sum X_i = t] = P(\sum_{i=1}^{n-2} X_i < X_{n-1} + X_n | \sum X_i = t)$$

$$= \frac{P(\sum_{i=1}^{n-2} X_i = t-1, X_n = 0, X_{n-1} = 1) + P(\sum_{i=1}^{n-2} X_i = t-1, X_n = 1, X_{n-1} = 0)}{P(\sum_{i=1}^n X_i = t)} +$$

$$\frac{P(\sum_{i=1}^{n-2} X_i = t, X_n = 0, X_{n-1} = 0) + P(\sum_{i=1}^{n-2} X_i = t-2, X_n = 1, X_{n-1} = 1)}{P(\sum_{i=1}^n X_i = t)}$$

$$= \frac{2\binom{n-2}{t-1}p^t(1-p)^{n-t} + \binom{n-2}{t}p^t(1-p)^{n-t} + \binom{n-2}{t-2}p^t(1-p)^{n-t}}{\binom{n}{t}p^t(1-p)^{n-t}}$$

$$= \frac{2\binom{n-2}{t-1} + \binom{n-2}{t} + \binom{n-2}{t-2}}{\binom{n}{t}}$$

By Rao-Blackwell theorem $F_1(\mathbf{X}) = \frac{2\binom{n-2}{\sum_{i=1}^n X_i - 1} + \binom{n-2}{\sum_{i=1}^n X_i} + \binom{n-2}{\sum_{i=1}^n X_i - 2}}{\binom{n}{\sum_{i=1}^n X_i}}$ is an unbiased estimator of $h(p_1)$. Since this estimator is a function of the complete sufficient statistics, it is the UMVUE for $h_1(p_1)$. Similarly, the UMVUE for $h_2(p_2)$ is $F_2(\mathbf{Y}) = \frac{2\binom{m-2}{\sum_{j=1}^m Y_j - 1} + \binom{m-2}{\sum_{j=1}^m Y_j} + \binom{m-2}{\sum_{j=1}^m Y_j - 2}}{\binom{m}{\sum_{j=1}^m Y_j}}$. Using the independence of $\mathbf{X}$ and $\mathbf{Y}$, we conclude that $F_1(\mathbf{X})F_2(\mathbf{Y})$ is the UMVUE for $h_1(p_1)h_2(p_2)$.

(c) The size of ϕ_1 is $E[\phi_1] = P(X_1 + X_2 + X_3 > 1) = \binom{3}{2}\left(\frac{1}{2}\right)^3 + \binom{3}{3}\left(\frac{1}{2}\right)^3 = \frac{1}{2}$.

The size of ϕ_2 is given by $E[\phi_2] = P(Y_1 + Y_2 + Y_3 + Y_4 > C) = \sum_{i=[C]+1}^{4} \binom{4}{i}\left(\frac{1}{3}\right)^i \left(\frac{2}{3}\right)^{4-i}$, where $[C]$ is the greatest integer smaller than C.

C can now be determined from the equation $\sum_{i=[C]+1}^{4} \binom{4}{i}\left(\frac{1}{3}\right)^i \left(\frac{2}{3}\right)^{4-i} < \frac{1}{2}$.

Problem 8 (AMS 206B):