בס"ד

התנהגות אסימפטוטית של שיטת נראות מותנית

$$\left.\begin{array}{l} X_1 \sim Bin(n_1, p_1) \\ X_2 \sim Bin(n_2, p_2) \end{array}\right\} \text{ב"ת} \qquad T = X_1 + X_2$$

Want to approximate

$$p(k|t) = Pr(X_1 = k \mid T = t)$$

Know this depends on p_1, p_2 only through $\psi = \frac{p_1 q_2}{p_2 q_1}$

So consider
$$\left.\begin{array}{l} \tilde{X}_1 \sim Bin(n_1, P_1) \\ \tilde{X}_2 \sim Bin(n_2, P_2) \end{array}\right\} \text{ב"ת} \qquad \tilde{T} = \tilde{X}_1 + \tilde{X}_2$$

where P_1, P_2 are the unique solution to

$$\frac{P_1 Q_2}{P_2 Q} = \psi, \qquad n_1 P_1 + n_2 P_2 = t$$

Then instead of looking at $Pr(X_1 = k \mid T = t)$ we can look at $Pr(\tilde{X}_1 = k \mid \tilde{T} = t)$.

Now, from the CLT we have

$$\left.\begin{array}{l} \tilde{X}_1 \mathrel{\dot\sim} N(n_1 P_1, n_1 P_1 Q_1) \\ \tilde{X}_2 \mathrel{\dot\sim} N(n_2 P_2, n_2 P_2 Q_2) \end{array}\right\} \text{ב"ת}$$

so that

$$\begin{bmatrix} \tilde{X}_1 \\ \tilde{T} \end{bmatrix} \mathrel{\dot\sim} N\left(\begin{bmatrix} n_1 P_1 \\ n_1 P_1 + n_2 P_2 \end{bmatrix}, \begin{bmatrix} n_1 P_1 Q_1 & n_1 P_1 Q_1 \\ n_1 P_1 Q_1 & n_1 P_1 Q_1 + n_2 P_2 Q_2 \end{bmatrix} \right)$$

By bivariate normal theory, we have that if

$$\begin{bmatrix} Y \\ Z \end{bmatrix} \sim N\left(\begin{bmatrix} \mu_Y \\ \mu_Z \end{bmatrix}, \begin{bmatrix} \sigma_Y^2 & \sigma_{YZ} \\ \sigma_{YZ} & \sigma_Z^2 \end{bmatrix} \right)$$

then the conditional dist'n of Y given Z is normal with

$$E[Y|Z] = \mu_Y + \frac{\sigma_{YZ}}{\sigma_Z^2}(Z - \mu_Z)$$

$$Var(Y|Z) = \sigma_Y^2 \left(1 - \frac{(\sigma_{YZ})^2}{\sigma_Y^2 \cdot \sigma_Z^2}\right)$$

Applying this to $\tilde{X}_1, \tilde{T}$ we get that the conditional dist'n of $\tilde{X}_1$ given $\tilde{T}$ is (approximately) normal with

$$\text{mean} = n_1 P_1 + \frac{n_1 P_1 Q_1}{n_1 P_1 Q_1 + n_2 P_2 Q_2}\left(t - [n_1 P_1 + n_2 P_2]\right)$$

$$= n_1 P_1$$

(since we defined P_1, P_2 to satisfy the condition $n_1 P_1 + n_2 P_2 = t$)

$$\text{variance} = n_1 P_1 Q_1 \left(1 - \frac{n_1 P_1 Q_1}{n_1 P_1 Q_1 + n_2 P_2 Q_2}\right)$$

$$= \frac{(n_1 P_1 Q_1)(n_2 P_2 Q_2)}{n_1 P_1 Q_1 + n_2 P_2 Q_2}$$

$$= \left(\frac{1}{n_1 P_1 Q_1} + \frac{1}{n_2 P_2 Q_2}\right)^{-1}$$

The approximate conditional means of all the table entries can be written as

$$\begin{array}{cc|c} E & n_1 - E & n_1 \\ t - E & n_2 - t + E & n_2 \\ \hline t & n_1 + n_2 - t & n_1 + n_2 \end{array}$$

with $E = n_1 P_1$. Now using $n_1 P_1 + n_2 P_2 = t$, we find that this is the same as

$$\begin{array}{cc|c} n_1 P_1 & n_1 Q_1 & n_1 \\ n_2 P_2 & n_2 Q_2 & n_2 \\ \hline t & n_1 + n_2 - t & n_1 + n_2 \end{array}$$

Thus we have

$$\frac{E(n_2 - t + E)}{(t - E)(n_1 - E)} = \frac{(n_1 P_1)(n_2 Q_2)}{(n_2 P_2)(n_1 Q_1)} = \frac{P_1 Q_2}{P_2 Q_1} = \psi$$

בס״ד

Conclusion: The conditional dist'n of X_1 given $T=t$ is approximately $N(E, V)$, where E satisfies the equation

$$\frac{E(n_2-t+E)}{(t-E)(n_1-E)} = \psi$$

and $$V = \left(\frac{1}{n_1P_1Q_1} + \frac{1}{n_2P_2Q_2}\right)^{-1}$$

where $$P_1 = \frac{E}{n_1}, \quad P_2 = \frac{t-E}{n_2}.$$