Example:

Sol: If $(Y_1, Y_2 \cdots Y_k) \sim$ Multinomial $(n: P_1, P_2, P_3 \cdots P_k)$

the probability mass fn' (PMF) is given by:

$$P(Y_1 = y_1, Y_2 = y_2 \cdots Y_k = y_k) = \frac{n!}{y_1! \, y_2! \cdots y_k!} P_1^{y_1} P_2^{y_2} \cdots P_k^{y_k}$$

$$Y_1 + Y_2 + Y_3 \cdots Y_k = n$$

Now, if $k = 2$

we have two categories with probabilities P_1 & P_2

$$P_1 + P_2 = 1$$

$$P(Y_1 = y_1, Y_2 = y_2) = \frac{n!}{y_1! \, y_2!} P_1^{y_1} P_2^{y_2}$$

$$Y_1 + Y_2 = n$$

$Y_2 = n - Y_1 \rightarrow$ knowing Y_1 completely determines Y_2

$$P(Y_1 = y_1) = \frac{n!}{y_1! \, (n - y_1)!} P_1^{y_1} P_2^{(n - y_1)}$$

$$= \frac{n!}{y_1! \, (n - y_1)!} P_1^{y_1} (1 - P_1)^{n - y_1}$$

Above expression is PMF of Binomial Random Variable.

$$Y_1 \sim \text{Binomial}(n, P_1)$$

3 Sol: $Y \sim \text{Binomial}(n, \pi)$, PMF is

$$P(Y=y) = \binom{n}{y} \pi^y (1-\pi)^{n-y}, \quad y = 0, 1, \dots n$$

The log-likelihood fn' is:

$$l(\pi) = \log \binom{n}{y} + y \log \pi + (n-y) \log(1-\pi)$$

To find MLE, we take derivative w.r.t π:

$$\frac{d}{d\pi} l(\pi) = \frac{y}{\pi} - \frac{n-y}{1-\pi}$$

to manimize it, set it equal to zero

$$\frac{d}{d\pi} l(\pi) = \frac{y}{\pi} - \frac{n-y}{1-\pi} = 0$$

$$\frac{y}{\pi} = \frac{n-y}{1-\pi}$$

$$\pi(n-y) = y(1-\pi)$$

$$\pi n - \pi y = y - \pi y$$

$$\boxed{\pi = y/n} \Rightarrow \hat{\pi} = \frac{y}{n}$$

Fisher Information: $I(\pi)$

$$I(\pi) = -E\left[\frac{d^2}{d\pi^2} l(\pi)\right]$$

$$\frac{d^2}{d\pi^2} l(\pi) = \frac{-y}{\pi^2} - \frac{n-y}{1-\pi^2}$$

Second derivative of log likelihood fn'

Considering expectations under $Y \sim \text{Bin}(n, \pi)$:

$$E[Y] = n\pi, \quad E[n-Y] = n(1-\pi)$$

so,

$$E\left[\frac{d^2}{d\pi^2} l(\pi)\right] = -\left(\frac{n\pi}{\pi^2} + \frac{n(1-\pi)}{(1-\pi)^2}\right)$$

$$= -\left(\frac{n}{\pi} + \frac{n}{1-\pi}\right)$$

$$= -n\left(\frac{1}{\pi(1-\pi)}\right)$$

$$I(\pi) = \frac{n}{\pi(1-\pi)}$$

The variance of $\hat{\pi} = \frac{Y}{n}$ is:

$$Var(\hat{\pi}) = Var\left(\frac{Y}{n}\right) = \frac{1}{n^2} Var(Y)$$

Since, $Y \sim Binomial(n, \pi)$ we know that:

$$Var(Y) = n\pi(1-\pi)$$

Thus,

$$Var(\hat{\pi}) = \frac{n\pi(1-\pi)}{n^2} = \frac{\pi(1-\pi)}{n}$$

Inverse Fisher info is:

$$I^{-}(\pi) = \frac{\pi(1-\pi)}{n}$$

Since, $Var(\hat{\pi}) = I^{-1}(\pi)$

Computing Score fn' $u(\pi; y)$

Score fn' is first derivative of log-likelihood fn'

$$u(\pi; y) = \frac{d}{d\pi} l(\pi)$$

$$u(\pi : y) = \frac{y}{\pi} - \frac{n-y}{1-\pi}$$

$$\boxed{u(\pi; y) = y/\pi - \frac{n-y}{1-\pi}}$$

4 Sol:

Rewriting likelihood fn'

$(Y_1, Y_2, Y_3 \ldots Y_k) \sim \text{Multinomial}(n; \pi_1, \pi_2, \pi_3 \ldots \pi_k)$ is

$$l(\pi_1, \pi_2, \ldots \pi_k \mid y_1, y_2 \ldots y_k) = \frac{n!}{y_1! \, y_2! \ldots y_k!} \pi_1^{y_1} \pi_2^{y_2} \ldots \pi_k^{y_k}$$

$$\pi_k = 1 - (\pi_1 + \pi_2 \ldots \pi_{k-1})$$

Now,

$$l(\pi_1, \pi_2 \ldots \pi_{k-1} \mid y_1, y_2, y_3 \ldots y_k) =$$

$$\frac{n!}{y_1! \, y_2! \ldots y_k!} \pi_1^{y_1} \pi_2^{y_2} \ldots \left(1 - \sum_{i=1}^{k-1} \pi_i\right)^{y_k}$$

Now, to find MLE's of $\pi_1, \pi_2 \ldots \pi_{k-1}$ we take log likelihood fn:

$$\log l(\pi_1, \pi_2 \ldots \pi_{k-1}) = \log \frac{n!}{y_1! y_2! \ldots y_k!} + y_1 \log \pi_1 + y_2 \log \pi_2 + \ldots$$

$$+ y_{k-1} \log \pi_{k-1} + y_k \log \left(1 - \sum_{i=1}^{k-1} \pi_i\right)$$

Taking partial derivaties w.r.t π_i for $i = 1 \ldots k-1$:

$$\frac{\delta}{\delta \pi_i} \log l = \frac{y_i}{\pi_i} - \frac{y_k}{1 - \sum_{j=1}^{k-1} \pi_j}$$

equating it to zero,

$$\frac{y_i}{\pi_i} = \frac{y_k}{1 - \sum_{j=1}^{k-1} \pi_j}$$

$$\pi_i = \frac{y_i}{y_k} \left(1 - \sum_{j=1}^{k-1} \pi_j\right)$$

Since,

$$\sum_{i=1}^{k} \pi_i = 1, \quad \pi_k = 1 - \sum_{j=1}^{k-1}$$

Substitute,

$$\sum_{i=1}^{k-1} \frac{y_i}{y_k} \left(1 - \sum_{j=1}^{k-1} \pi_j\right) + \left(1 - \sum_{i=1}^{k-1} \pi_i\right) = 1$$

Thus,

$$\sum_{i=1}^{k} \pi_i = 1$$

Using method of Lagrange ~~equation~~ multipliers..

$$\hat{\pi}_i = \frac{y_i}{n} \quad \text{for } i = 1, \ldots k$$

Thus, MLEs of π_i are simply observed proportions $\hat{\pi}_i = \frac{y_i}{n}$ $i = 1, \ldots k$

Example 5:

$$L(\pi|y) = \binom{n}{y}\pi^{y}(1-\pi)^{n-y}$$

$$\log L(\pi|y) = y\log\pi + (n-y)\log(1-\pi) + c$$

Score fn'

It is first derivative of log-likelihood fn'

$$u(\pi) = \frac{d}{d\pi}\log L(\pi)$$

$$= \frac{y}{\pi} - \frac{(n-y)}{1-\pi}$$

Setting $\pi = \pi_0$, under H_0, Score test statistic is

$$Z_S = \frac{\hat{\pi} - \pi_0}{\sqrt{\frac{\pi_0(1-\pi_0)}{n}}}$$

$$Z_S = \frac{u(\pi_0)}{\sqrt{I(\pi_0)}}$$

$$I(\pi_0) = \frac{d}{d\pi}\log L(\pi)$$

$$= \frac{n}{\pi_0(1-\pi_0)}.$$

$$Z_S = \frac{y/\pi_0 - \frac{(n-y)}{1-\pi_0}}{\sqrt{\frac{n}{\pi_0(1-\pi_0)}}}$$

$\hat{\pi} = Y/n$.

$$Z_s = \frac{\hat{\pi} - \pi_0}{\sqrt{\frac{\pi_0(1-\pi_0)}{n}}}$$

$$Z_s = \frac{\hat{\pi} - \pi_0}{\sqrt{\frac{\pi_0(1-\pi_0)}{n}}}$$

Ex 5:

Likelihood ratio:

The likelihood fn under null hypothesis H_0: $\pi = \pi_0$ is

$$L(\pi_0) = \binom{n}{y} \pi_0^y (1-\pi_0)^{n-y}$$

MLE is $\hat{\pi} = \frac{Y}{n}$

likelihood of MLE is

$$L(\hat{\pi}) = \binom{n}{y} \hat{\pi}^y (1-\hat{\pi})^{n-y}$$

likelihood Ratio is

$$\Lambda = \frac{L(\pi_0)}{L(\hat{\pi})} = \frac{\pi_0^y (1-\pi_0)^{n-y}}{\hat{\pi}^y (1-\hat{\pi})^{n-y}}$$

log likelihood Ratio

$$\ln \Lambda = y \ln \frac{\pi_0}{\hat{\pi}} + (n-y) \ln \frac{1-\pi_0}{1-\hat{\pi}}$$

Multiplying with -2;

$$G^2 = -2 \ln \Lambda$$

$$G^2 = 2\left[y \ln \frac{y}{n\pi_0} + (n-y) \ln \frac{n-y}{n(1-\pi_0)}\right]$$

n

<u>Example 6:</u>

Given,

$$Y \sim \text{Binom}(n=100, \pi)$$

$H_0 : \pi = 0.5$

with $n = 100$, $Y = 60$.

$$\hat{\pi} = \frac{Y}{n} = \frac{60}{100} = 0.6$$

<u>Wald Test:</u>

$$Z_W = \frac{\hat{\pi} - \pi_0}{\sqrt{\frac{\hat{\pi}(1-\hat{\pi})}{n}}}$$

$$= \frac{0.6 - 0.5}{\sqrt{\frac{0.6(1-0.6)}{100}}} = \frac{0.1}{\sqrt{0.002}} = 2.04$$

P-value – $p = 0.0414$

$p < 0.05$, reject H_0

Score Test Statistic:

$$z_s = \frac{\hat{\pi} - \pi_0}{\sqrt{\frac{\pi_0(1-\pi_0)}{n}}}$$

$$= \frac{0.6 - 0.5}{\sqrt{\frac{0.5(1-0.5)}{100}}} = 2.0$$

$$p = 2(1 - P(z \le 2.0))$$

$$= 2(1 - 0.9772) = 0.0456$$

$p < 0.05$, reject H_0

Likelihood Ratio Test Statistic:

$$-2\log\Lambda = 2\left[y\log\frac{y}{n\pi_0} + (n-y)\log\frac{n-y}{n(1-\pi_0)}\right]$$

$$= 2\left[60\log\frac{60}{100 \times 0.5} + 40\log\frac{40}{100 \times 0.5}\right]$$

$$= 2\left[60\log 1.2 + 40\log 0.8\right]$$

$$= 2\left[60(0.1823) + 40(\cancel{-0.0969}\ -0.223)\right]$$

$$= \cancel{14.124}\ 4$$

Chi square table: $\alpha = 0.005$ $\chi^2_{0.05} = 3.84$)

~~$p = P(\chi^2_1 \ge 14.124)$~~ ~~$\approx$~~ 0.0002

$4 > 3.84$)

~~$p < 0.05$, reject H_0.~~ Reject H_0 //.

Sol: Given,

$n = 20$ (trails)

$y = 0$ (new drug is worse each time)

$$\hat{\pi} = \frac{y}{n} = \frac{0}{20} = 0$$

a) 95% Wald confidence interval

$$\hat{\pi} \pm z_{\alpha/2} \sqrt{\frac{\hat{\pi}(1-\hat{\pi})}{n}}$$

$$= 0 \pm 1.96 \times \sqrt{\frac{0(1-0)}{20}}$$

$$= (0, 0)$$

↳ Wald CI fails because it doesn't account for boundary cases.

b) 95% adjusted Wald (Agresti-Coull)

$$\tilde{n} = n + z_{\alpha/2}^2 = 20 + 1.96^2 = 20 + 3.84 = 23.84$$

$$\tilde{\pi} = \frac{y + z_{\alpha/2}^2/2}{n + z_{\alpha/2}^2} = \frac{0 + 3.84/2}{23.84} = \frac{1.92}{23.84}$$

$$= 0.0806$$

SE:

$$= \sqrt{\frac{\tilde{\pi}(1-\tilde{\pi})}{\tilde{n}}} = \sqrt{\frac{(0.0806)(1-0.0806)}{23.84}}$$

$$\simeq 0.0596$$

Confidence interval:

$$\tilde{\pi} \pm z_{\alpha/2}\, SE$$

$0.0806 \pm 1.96 \times 0.0596$

0.0806 ± 0.117

$(0, 0.198)$

c) 95% Score (wilson) Confidence Interval

$$\frac{\hat{\pi} \pm \frac{z^2_{\alpha/2}}{2n} + z_{\alpha/2}\sqrt{\frac{\hat{\pi}(1-\hat{\pi})}{n} + \frac{z^2_{\alpha/2}}{4n^2}}}{1 + \frac{z^2_{\alpha/2}}{n}}$$

$$\frac{0 + \frac{3.84}{40} \pm 1.96\sqrt{\frac{3.84}{1600}}}{1 + \frac{3.84}{20}}$$

$$\Rightarrow \frac{0.096 \pm 1.96 \times 0.049}{1.192}$$

$$= \frac{0.096 \pm 0.096}{1.192}$$

$(0, 0.161)$

d) 95% likelihood Ratio Confidence interval

$$-2\ln\Lambda \leq \chi^2_{1,0.05} = 3.841$$

likelihood ratio for binomial is:

$$l(\pi) = \binom{20}{0}\pi^0(1-\pi)^{20}$$

Maximizing $l(\pi)$ gives $\hat{\pi} = 0$, so LR test statistic

$$40\ln\frac{1}{1-\pi} = 3.841$$

$$\ln\frac{1}{1-\pi} = 0.096$$

$$\frac{1}{1-\pi} = e^{0.096} \approx 1.101$$

$$1-\pi = \frac{1}{1.101} \approx 0.908$$

$$\boxed{\pi = 0.092}$$

$\therefore$ likelihood Ratio CI is $(0, 0.092)$

8 Sol: To prove

$$\chi^2 = Z_s^2 \quad \text{when } k=2$$

$$G^2 = -2\ln\Lambda = 2\sum_{j=1}^{k} y_j \ln\left(\frac{y_i}{n\pi_j^0}\right)$$

$\Rightarrow$

$$Z_s = \frac{\hat{\pi} - \pi_0}{\sqrt{\frac{\pi_0(1-\pi_0)}{n}}}$$

$$Z_s^2 = \frac{n(\hat{\pi} - \pi_0)^2}{\pi_0(1-\pi_0)}$$

Chi square for $k=2$

$$\chi^2 = \sum_{j=1}^{2} \frac{(y_i - n\pi_j^0)^2}{n\pi_j^0}$$

$y_1 = n\hat{\pi} \quad y_2 = n(1-\hat{\pi})$

$$\chi^2 = \frac{n(\hat{\pi} - \pi_0)^2}{\pi_0(1-\pi_0)}$$

$$\therefore \boxed{Z_s^2 = \chi^2}$$

Likelihood for multinomial distribution;

$$L(\pi_1, \pi_2 \cdots \pi_k) = \frac{n!}{y_1!\, y_2! \cdots y_k!} \pi_1^{y_1} \pi_2^{y_2} \pi_3^{y_3} \cdots y_k^{y_k}$$

likelihood at π_0

$$G^2 = -2\ln\Lambda = -2\ln\frac{L(\pi_0)}{L(\hat{\pi})}$$

$$G^2 = 2\sum_{j=1}^{k} y_j \ln\left(\frac{y_j}{n\pi_j^0}\right)$$

$$\Lambda = \frac{L(\pi_0)}{L(\hat{\pi})} = \frac{\prod_{j=1}^{k} (\pi_j^0)^{y_j}}{\prod_{j=1}^{k} \left(\frac{y_j}{n}\right)^{y_j}}$$

taking log

$$\ln \Lambda = \sum_{j=1}^{k} y_j \ln \pi_j^0 - \sum_{j=1}^{k} y_j \ln \frac{y_j}{n}$$

$$\ln \Lambda = \sum_{j=1}^{k} y_j \ln\left(\frac{\pi_j^0}{y_j/n}\right)$$

$$G^2 = -2 \ln \Lambda = 2\sum_{j=1}^{k} y_i \ln\left(\frac{y_j}{n\pi_j^0}\right)$$

//.