1"02

Suissa + Shuster (1985, JRSS A)

Exact Size of Z-Test

$$Z = \frac{\hat{p}_1 - \hat{p}_0}{\left[\left(\frac{1}{n_{0\cdot}} + \frac{1}{n_{1\cdot}}\right) \hat{p}(1-\hat{p})\right]^{1/2}}$$

$$Z \geq z_\alpha$$

$$\iff (n_{01}, n_{11}) \in C$$

for some region C in $\mathbb{R}^2$.

The region C can be characterized appropriately with some elementary algebra

Under $H_0: p_0 = p_1 = p$, we have

$$\Pi(p) = \Pr(Z \geq z_\alpha)$$

$$= \sum_{(n_{01}, n_{11}) \in C} \binom{n_{0\cdot}}{n_{01}} \binom{n_{1\cdot}}{n_{11}} p^{n_{01}+n_{11}} (1-p)^{n_{\cdot\cdot} - (n_{01}+n_{11})}$$

Note that this depends on p, which is unknown.

We can define the p-value as $\Pi(p^*)$, where $p^* = \arg\max \Pi(p)$.

p^* can be found by a search routine, eg. Matlab fminbnd

The algebra:

Put $\hat{\Delta} = \hat{p}_1 - \hat{p}_0$

Then $Z \geq z_\alpha$ is equivalent to

$$\hat{\Delta} \geq z_\alpha \left[\left(\frac{1}{n_{0\cdot}} + \frac{1}{n_{1\cdot}} \right) \hat{\bar{p}} (1-\hat{\bar{p}}) \right]^{1/2} \qquad (*)$$

Now,

$$\hat{\bar{p}} = \frac{n_{0\cdot}\hat{p}_0 + n_{1\cdot}\hat{p}_1}{n_{\cdot\cdot}}$$

$$= \hat{p}_0 + \frac{n_{1\cdot}}{n_{\cdot\cdot}} \hat{\Delta}$$

$$= \hat{p}_0 + \omega\hat{\Delta}, \qquad \omega = \frac{n_{1\cdot}}{n_{\cdot\cdot}}$$

$$1 - \hat{\bar{p}} = \hat{q}_0 - \omega\hat{\Delta}, \qquad \hat{q}_0 = 1 - \hat{p}_0$$

$$\hat{\bar{p}}(1-\hat{\bar{p}}) = (\hat{p}_0 + \omega\hat{\Delta})(\hat{q}_0 - \omega\hat{\Delta})$$

$$= \hat{p}_0\hat{q}_0 + \omega(\hat{q}_0 - \hat{p}_0)\hat{\Delta} - \omega^2\hat{\Delta}^2$$

So, the condition $(*)$ is equivalent to $\Delta \geq \Delta^*$, where Δ^* is the relevant root of the quadratic eqn

$$\Delta^2 = z_\alpha^2 \left(\frac{1}{n_{0\cdot}} + \frac{1}{n_{1\cdot}}\right) \left[\hat{p}_0\hat{q}_0 + \omega(\hat{q}_0 - \hat{p}_0) - \omega^2\Delta^2 \right]$$

i.e. $$A\Delta^2 + B\Delta + C = 0$$

with $$A = 1 + z_\alpha^2 \left(\frac{1}{n_{0\cdot}} + \frac{1}{n_{1\cdot}}\right)\omega^2$$

$$B = z_\alpha^2 \left(\frac{1}{n_{0\cdot}} + \frac{1}{n_{1\cdot}}\right)\omega(\hat{p}_0 - \hat{q}_0)$$

$$C = -z_\alpha^2 \left(\frac{1}{n_{0\cdot}} + \frac{1}{n_{\cdot}}\right)\hat{p}_0\hat{q}_0$$

That is,

$$\Delta^* = \frac{-B + \sqrt{B^2 - 4AC}}{2A}$$

Thus:

$$(*) \iff \hat{\Delta} \geq \Delta^*$$

$$\iff \hat{p}_1 \geq p_0 + \Delta^*$$

$$\iff n_{11} \geq n_{1\cdot}(\hat{p}_0 + \delta^*)$$

The quantity $n_{1\cdot}(\hat{p}_0 + \Delta^*)$ depends on n_{01}, so we will write it as $h(n_{01})$.

We then have

$$\Pi(p) = \sum_{m=0}^{n_{0\cdot}} b(n_{0\cdot}, p, m)\left(1 - B(n_{1\cdot}, p, h(m))\right)$$