$$\chi^2_{(r-1)\times(c-1)} = \sum_{i=1} \frac{(O_i - e_i)^2}{e_i} \quad 即 \sum_{i=1} \frac{(真实值i - 计算值i)^2}{计算值i = n \times \mathbb{P}_{row} \times \mathbb{P}_{col}}$$

$$\text{P-value} = \underbrace{\mathbb{P}(\chi^2_{(r-1)(c-1)} \geq \chi^2)}_{写这么写} \text{ is between \_\_ and \_\_}$$

首先对表进行假设

H_0: row and col are independent

H_1: some association between row and col.

$$\mathbb{P}\text{-value} = \mathbb{P}\left(Z_{\frac{\alpha}{2}} / t_{\frac{\alpha}{2}}(n-1) / \chi^2_{(r-1)(c-1)} \geq \frac{\bar{x}-\mu}{\sqrt{\frac{\sigma^2}{n}}} \Big/ \frac{\bar{x}-\mu}{\sqrt{\frac{s^2}{n}}} \Big/ \sum_{i=0} \frac{(O_i-e_i)^2}{e_i}\right)$$

$$\text{P-value} \Rightarrow Z_\alpha \geq \frac{\bar{P}_1 - \bar{P}_2 - (\mu_1 - \mu_2)}{\sqrt{\frac{\bar{P}(1-\bar{P})}{n_1} + \frac{\bar{P}(1-\bar{P})}{n_2}}}$$

P值分母$\bar{P}$为总数平均

$$\text{a margin of error} \geq Z_{\frac{\alpha}{2}} \times \sqrt{\frac{S_p^2}{n}} \Rightarrow n$$

n (minimum) $|\leq|$ within a margin of error $|\geq|$

$y = \underline{m}x + \underline{c}$

slope β_1, intercept β_0

Slope: One ~~additional~~ additional X:name is associated with an increase / decrease of m in Y:name

Intercept: For some with no X:name, we expect a Y:name of c

$R^2 \sim$ R-squared.

only R-squared of the variation in Y:name is explained by the X:name (then biger the better) (0,1).

Root Mean Squared Error: The standard deviation of fluctuation about the regression line is estimated as 值