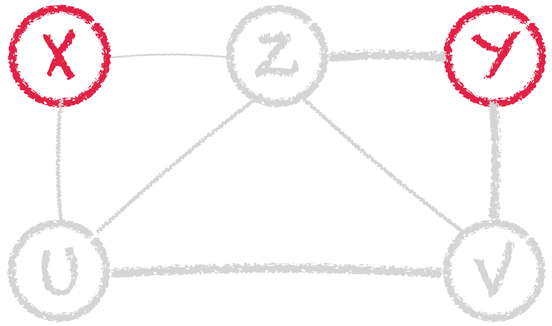
divergence tests of goodness of fit

testing pairwise independence

of random variable X and Y with r_X and r_Y outcomes

$p =$ model based on empirical distribution $p(x, y)$ with $d(p) = r_X r_Y - 1$

$q = X \perp Y$ such that $p(x) \cdot p(y)$ with $d(q) = (r_X - 1) + (r_Y - 1)$



☑ log likelihood ratio test statistic

$$\begin{aligned}\chi^2((r_X - 1)(r_Y - 1)) &= 2nD(p, q) \\ &= 2n[H(X) + H(Y) - H(X, Y)] \\ &= 2nJ(X, Y)\end{aligned}$$

☑ independence is rejected if

$$\chi^2((r_X - 1)(r_Y - 1)) \geq (r_X - 1)(r_Y - 1) + \sqrt{8(r_X - 1)(r_Y - 1)}$$

or if the empirical joint entropy $J(X, Y)$ is larger than $[(r_X - 1)(r_Y - 1) + \sqrt{8(r_X - 1)(r_Y - 1)}]/2n$

divergence tests of goodness of fit

testing conditional independence

of random variable X, Y and Z with r_X, r_Y and r_Z outcomes

$p =$ model based on empirical distribution $p(x, y, z)$ with $d(p) = r_X r_Y r_Z - 1$

$q = X \perp Y | Z$ such that $p(x, z)p(y, z)/p(z)$ with $d(q) = r_Z - 1 + r_Z(r_X - 1 + r_Y - 1)$