divergence tests of goodness of fit

testing nested model specifications

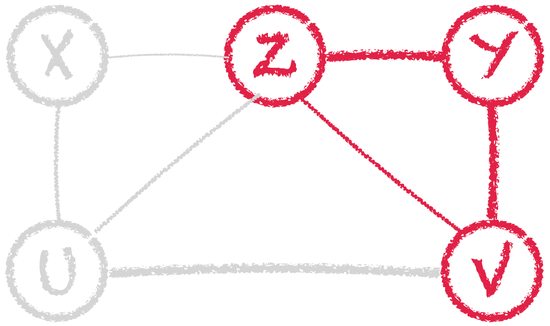
example: five dimensional data (X, Y, Z, U, V) with r_X, r_Y, r_Z, r_U, r_V outcomes

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

$q =$ model with listed imposed independence and conditional independence assumptions

examples: $q_1 = X \perp (Y, Z, U, V)$ and $U \perp (Y, Z, V)$

$q_2 = X \perp (Y, Z, U, V)$ and $U \perp (Y, Z, V)$ and $Z \perp V \mid Y$



divergence tests of goodness of fit

testing nested model specifications

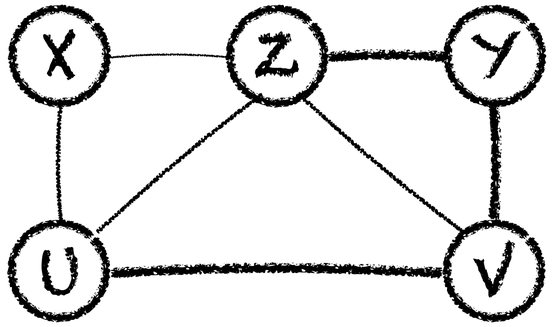
example: five dimensional data (X, Y, Z, U, V) with r_X, r_Y, r_Z, r_U, r_V outcomes

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

$q =$ model with listed imposed independence and conditional independence assumptions

examples: $q_1 = X \perp (Y, Z, U, V)$ and $U \perp (Y, Z, V)$

$q_2 = X \perp (Y, Z, U, V)$ and $U \perp (Y, Z, V)$ and $Z \perp V \,|\, Y$



☑ $d(q)$ is sum of the degrees of freedom of its independent components

☑ divergence $D(p, q)$ of each model is the sums of the divergences of its nested specifications

$$\chi^2(d) = 2nD(p, q_1) = 2n[D(X \perp (Y, Z, U, V)) + D(U \perp Y, Z, V)]$$

$$\chi^2(d) = 2nD(p, q_2) = 2n[D(X \perp (Y, Z, U, V)) + D(U \perp Y, Z, V) + D(Z \perp V \,|\, Y)]$$

where d can be obtained as either

✓ the sums of degrees of freedom for the divergences of the nested specifications

✓ the difference between degrees of freedom of the general and the specified model $d(p) - d(q)$