- Variable types: Nominal/Ordinal (Qualitative); Interval/Ratio (Quantitative); Continuous/Discrete
- Binomial: $p(y) = \binom{n}{y}\pi^y(1-\pi)^{n-y}, y = 0, \ldots, n$
  $\mathsf{E}(Y_i) = \pi, \quad \mathsf{Var}(Y_i) = \pi(1-\pi)$
  $\mu = \mathsf{E}(Y) = n\pi, \quad \sigma^2 = \mathsf{Var}(Y) = n\pi(1-\pi)$
  ML: $\hat\pi = y/n, \quad \mathsf{E}(\hat\pi) = \pi, \quad \sigma(\hat\pi) = \sqrt{\frac{\pi(1-\pi)}{n}}$
  Wald: $z_W = \dfrac{\hat\pi - \pi_0}{\sqrt{\hat\pi(1-\hat\pi)/n}}$, CI: $\hat\pi \pm z_{\alpha/2}\sqrt{\dfrac{\hat\pi(1-\hat\pi)}{n}}$
  Score: $z_S = \dfrac{\hat\pi - \pi_0}{\sqrt{\pi_0(1-\pi_0)/n}}$
- Multinomial: $p(n_1, .., n_{c-1}) = \left(\frac{n!}{n_1!\cdots n_c!}\right)\pi_1^{n_1}\cdots\pi_c^{n_c}$
  $\mathsf{E}(n_j) = n\pi_j, \quad \mathsf{Var}(n_j) = n\pi_j(1-\pi_j)$
  $\mathsf{Cov}(n_j, n_k) = -n\pi_j\pi_k$, ML estimate: $\hat\pi_j = n_j/n$
- Poisson: $p(y) = \dfrac{e^{-\mu}\mu^y}{y!}, \quad \mathsf{E}(Y) = \mathsf{Var}(Y) = \mu$
- $E(Y) = \mathsf{E}[\mathsf{E}(Y|\mu)], \mathsf{Var}(Y) = \mathsf{E}[\mathsf{Var}(Y|\mu)] + \mathsf{Var}[\mathsf{E}(Y|\mu)]$
- Information matrix (j,k) element: $-\mathsf{E}\left(\dfrac{\partial^2 L(\beta)}{\partial\beta_j\partial\beta_k}\right)$
- Standard errors are square roots of diagonal elements of inverse information matrix.
- Information: $\iota(\theta) = \mathsf{E}[-L''(\theta)]$, $\hat V(\hat\theta) = i^{-1}(\theta)|_{\theta=\hat\theta}$
- Wald test stat. $z = (\hat\beta - \beta_0)/SE, CI : \hat\beta \pm z_{\alpha/2}SE$
- Likelihood-ratio test stat.: $\chi^2_{\dim(H_a \cup H_0) - \dim(H_0)}$
  $-2\log(\Lambda) = -2\log(\ell_0/\ell_1) = -2(L_0 - L_1)$
- Score function: $u(\beta) = \partial L(\beta)/\partial\beta$
- Chi-squared form of score statistic:
  $$\frac{[u(\beta_0)]^2}{\iota(\beta_0)} = \frac{[\partial L(\beta)/\partial\beta_0]^2}{-\mathsf{E}\left[\partial^2 L(\beta)/\partial\beta_0^2\right]} = \frac{\text{score}^2}{\text{information}} \sim \chi_1^2$$
- Pearson test for specific multinomial $H_0 : \pi_j = \pi_{j0}$:
  $$X^2 = \sum_j \frac{(n_j - \mu_j)^2}{\mu_j} \sim \chi^2_{c-1}, \quad \mu_j = n\pi_{j0}$$
- Likelihood-ratio chi-squared statistic:
  $$G^2 = -2\log\Lambda = 2\sum n_j \log(n_j/n\pi_{j0}) \sim \chi^2_{c-1}$$

Row	1	2	Marginal
1	π_{11}	π_{12}	$\pi_{1+} = \sum_j \pi_{1j}$
	$(\pi_{1\|1})$	$(\pi_{2\|1})$	(1.0)
1	π_{21}	π_{22}	$\pi_{2+} = \sum_j \pi_{2j}$
	$(\pi_{1\|2})$	$(\pi_{2\|2})$	(1.0)
Marg.	$\pi_{+1} = \sum_i \pi_{i1}$	$\pi_{+2} = \sum_i \pi_{i2}$	1.0

- Sensitivity: p(test positive | has disease) $= \pi_{1|1}$
- Specificity: p(test negative | no disease) $= \pi_{2|2}$
- Conditional: $\pi_{j|i} = \pi_{ij}/\pi_{i+}$ for all i and j
- Independent: $\pi_{ij} = \pi_{i+}\pi_{+j}$ for all i and j
- Homogeneity: X explanatory, $\pi_{j|1} = \cdots = \pi_{j|I}$
- Difference of proportions: $\pi_1 - \pi_2$ $(\pi_i \equiv \pi_{1|i})$
- Relative risk: π_1/π_2 ($1 \equiv$ independence)
- Odds: $\Omega = \pi/(1-\pi)$ $(\pi = \Omega/(\Omega+1))$
- Odds ratio: $\theta = \dfrac{\Omega_1}{\Omega_2} = \dfrac{\pi_1/(1-\pi_1)}{\pi_2/(1-\pi_2)} = \dfrac{\pi_{11}\pi_{22}}{\pi_{12}\pi_{21}}$
  odds ratio = relative risk $\left(\frac{1-\pi_2}{1-\pi_1}\right)$
- Sample odds ratio: $\hat\theta = n_{11}n_{22}/n_{12}n_{21}$
- Partial table: XY contingency table for fixed Z
- XY marginal table: sums partial tables over Z
- $2 \times 2 \times K$ tables, $K \equiv \#$ of control categories:
  $\{\mu_{ijk}\} \equiv$ cell expected frequencies
- XY conditional odds ratio: $\theta_{XY(k)} = \dfrac{\mu_{11k}\mu_{22k}}{\mu_{12k}\mu_{21k}}$
- XY marginal table expected frequencies:
  $\{\mu_{ij+} = \sum_k \mu_{ijk}\}$
- XY marginal odds ratios: $\theta_{XY} = \dfrac{\mu_{11+}\mu_{22+}}{\mu_{12+}\mu_{21+}}$
- Sample counts use cell counts for $\hat\theta_{XY(k)}$ and $\hat\theta_{XY}$.
- Conditionally independent at level k of Z:
  $P(Y = j|X = i, Z = k) = P(Y = j|Z = k)\forall i, j$
- Conditional independence given Z:
  $\pi_{ijk} = \pi_{i+k}\pi_{+jk}/\pi_{++k}$ for all i, j, k
- Homogeneous XY assoc: $\theta_{XY(1)} = \cdots = \theta_{XY(k)}$
- Mid-P-value: $\frac{1}{2}P(t = t_o) + P(T > t_o)$

- Sample odds ratio: $\mathsf{E}(\hat\theta)$ and $\mathsf{Var}(\hat\theta)$ undefined.
  Log odds ratio CI: $\log\hat\theta \pm z_{\alpha/2}\hat\sigma(\log\hat\theta)$
  $\hat\sigma(\log\hat\theta) = \left(\frac{1}{n_{11}} + \frac{1}{n_{12}} + \frac{1}{n_{21}} + \frac{1}{n_{22}}\right)^{1/2}$
- Diff. of proportions: $\mathsf{E}(\hat\pi_1 - \hat\pi_2) = \hat\pi_1 - \hat\pi_2$
  Wald CI: $(\hat\pi_1 - \hat\pi_2) \pm z_{\alpha/2}\hat\sigma(\hat\pi_1 - \hat\pi_2)$
  $\hat\sigma(\hat\pi_1 - \hat\pi_2) = \left[\frac{\hat\pi_1(1-\hat\pi_1)}{n_1} + \frac{\hat\pi_2(1-\hat\pi_2)}{n_2}\right]^{1/2}$
- Sample relative risk: $r = \hat\pi_1/\hat\pi_2$
  Wald CI: $\log r \pm z_{\alpha/2}\hat\sigma(\log r)$
  $\sigma(\log r) = \left(\frac{1-\pi_1}{\pi_1 n_1} + \frac{1-\pi_2}{\pi_2 n_2}\right)^{1/2}$
- Delta method: $T_n \sim N(\theta, \sigma/\sqrt{n})$, $g(\theta)$:
  Wald CI: $g(T_n) \pm z_{\alpha/2}|g'(T_n)|\sigma(T_n)\sqrt{n}$
- Pearson chi-squared test of independence:
  $X^2 = \sum_i \sum_j \frac{(n_{ij} - \hat\mu_{ij})^2}{\hat\mu_{ij}} \sim \chi^2_{(I-1)(J-1)}$
  Est. expected frequencies: $\hat\mu_{ij} = n_{i+}n_{+j}/n$
- Likelihood-ratio chi-squared test:
  $G^2 = -2\log\Lambda = 2\sum_i\sum_j n_{ij}\log(n_{ij}/\hat\mu_{ij})$
- Pearson residual: $e_{ij} = \frac{n_{ij} - \hat\mu_{ij}}{\hat\mu^{1/2}}$
- Standardized Pearson residual:
  $\frac{n_{ij} - \hat\mu_{ij}}{[\hat\mu_{ij}(1-p_{+i})(1-p_{+j})]^{1/2}} \sim N(0,1)$
- $I \times J$ table chi-squared partition:
  $$\frac{\sum_{a<i}\sum_{b<j} n_{ab} \quad | \quad \sum_{a<i} n_{aj}}{\sum_{b<j} n_{ib} \quad | \quad n_{ij}}$$
- Ordinal: $M^2 = (n-1)r^2$
- $\gamma = \frac{\Pi_c - \Pi_d}{\Pi_c + \Pi_d}, \quad \hat\gamma = \frac{C-D}{C+D}$
  $\Pi_c = 2\sum_i\sum_j \pi_{ij}\left(\sum_{h>i}\sum_{k>j}\pi_{hk}\right)$
  $\Pi_d = 2\sum_i\sum_j \pi_{ij}\left(\sum_{h>i}\sum_{k<j}\pi_{hk}\right)$
- Fisher's exact test generates all tables consistent with given margin totals: $H_a : \theta > 1$
  p-value: $P(n_{11} \geq t_o)$, $t_o \equiv$ observed n_{11}
  $p(t) = P(n_{11} = t) = \dfrac{\binom{n_{1+}}{t}\binom{n_{2+}}{n_{+1}-t}}{\binom{n}{n_{+1}}}$