# MedicalBiostatistics.com

## TESTS OF SIGNIFICANCE

**Table 1:** Statistical tests of significance for qualitative data (proportions)

| Parameter of Interest and Setup | Conditions | Main Criterion for Test of Significance |
|---|---|---|
| **Proportions in all of the Following Small Sized Tables** | | |
| One dichotomous variable | Independent trials | |
| | Any *n* | Binomial |
| | Large *n* | Gaussian *Z* |
| One polytomous variable | Independent trials | |
| | Large *n* | Goodness-of-fit chi-square |
| | Small *n* | Multinomial |
| Two dichotomous variables (2×2) | Two independent samples | |
| | Large *n* | Chi-square or Gaussian *Z* |
| | Small *n* | Fisher exact |
| | Detecting a medically important difference—Large *n* | Gaussian *Z* |
| | Equivalence test | TOSTs |
| | Matched pairs | |
| | Large *n* | McNemar |
| | Small *n* | Binomial |
| | Crossover design | |
| | Large *n* | Chi-square |
| | Small *n* | Fisher exact |
| **Bigger Tables, No Matching** | | |
| Association | 2×*C* tables – large *n* | Chi-square |
| Trend in proportions | 2×*C* tables – large *n* | Chi-square for trend |
| Dichotomy in repeated measures | Many related 2×2 tables | Cochran *Q* |
| Association | *R*×*C* tables | |
| Association | Three-way tables | Chi-square |
| | Test of full independence | Chi-square |
| | Test of other types of independence (log-linear models) | $G^2$ |
| ***I*×*I* Table** | Matched pairs | McNemar–Bowker |
| **Association in Stratified Tables** | Stratified into many 2×2 tables | Mantel-Haenszel chi-square |

**Table 2:** Statistical tests of significance for relative risk (RR) and odds ratio (OR)

| Parameter of Interest and Setup | Conditions | Main Criterion for Test of Significance |
|---|---|---|
| **Relative Risk, Odds Ratio, and Attributable Risks** | **Large *n* required** | |
| ln(RR) or ln(OR) | One group | Wald |
| ln(RR) or ln)OR) | Two independent samples | Gaussian *Z* or Chi-square |
| RR or OR | Matched pairs | Gaussian *Z* or McNemar |

| | | |
|---|---|---|
| | Stratified | Mantel–Haenszel chi-square |
| AR | Two independent samples | Chi-square or Gaussian $Z$ |
| | Matched pairs | McNemar |
| Homogeneity of RRs or ORs across strata | Large sample | Breslow–Day |
| | Small sample | Zelen test |

**Table 3:** Statistical tests of significance for quantitative data (means, variances, correlations, survival))

| Parameter of Interest and Setup | Conditions | Main Criterion for Test of Significance |
|---|---|---|
| **Mean or Central value** | | |
| One group | Comparison with prespecified – Gaussian | |
| | $\sigma$ known | Gaussian $Z$ |
| | $\sigma$ not known | Student $t$ |
| Comparison of two independent groups | Paired – Gaussian ($\sigma$ not known) | Student $t$ |
| | Paired – NonGaussian | |
| | Any $n$ | Sign test |
| | $5 \leq n \leq 19$ | Wilcoxon signed-ranks $W_S$ |
| | $20 \leq n \leq 29$ | Standardized $W_S$ referred to Gaussian Z |
| | $n \geq 30$ | Student $t$ |
| | Unpaired – Gaussian | |
| | Equal variances | Student $t$ |
| | Unequal variances | Welch |
| | Unpaired – NonGaussian | |
| | $n_1$, $n_2$ between (4, 9) | Wilcoxon rank-sum $W_R$ |
| | $n_1$, $n_2$ between (10, 29) | Standardized $W_R$ referred to Gaussian Z |
| | $n_1$, $n_2 \geq 30$ | Student $t$ |
| | Crossover design – Gaussian | Student $t$ |
| | Detecting medically important difference | Student $t$ |
| | Equivalence tests | Student $t$ (TOSTs) |
| Comparison of three or more independent groups, equal variances | One-way, two-way or multi-way layout – Gaussian | ANOVA $F$ |
| | One-way Nonparametric | |
| | $n \leq 5$ | Kruskal–Wallis $H$ |
| | $n \geq 6$ | $H$ referred to chi-square |
| | Two-way layout – Gaussian | ANOVA $F$ |
| | Two-way Nonparametric (one observation per cell) | |
| | $J \leq 13$ and $K = 3$ | Friedman $S$ |
| | $J \leq 8$ and $K = 4$ | Friedman $S$ |
| | $J \leq 5$ and $K = 5$ | Friedman $S$ |
| | Larger $J$, $K$ | $S$ referred to chi-square |
| | Multiple comparisons – Gaussian | |
| | All pairwise | Tukey $D$ |
| | With control group | Dunnett |
| | Few comparisons | Bonferroni |
| Repeated measures | Gaussian | ANOVA $F$ with Hyun–Feldt correction for sphericity |
| Comparison of three or more groups, unequal variances | Gaussian | |
| | Large samples | Welch |

| | | |
|---|---|---|
| | Small samples | Brown–Forsyth |
| **Variance** | | |
| One group | Comparison with prespecified – | |
| | Gaussian | Variance ratio F |
| Two independent groups | Gaussian | Variance ratio F |
| | NonGaussian (Mild) | Levene |
| More than two independent groups | Gaussian | Bartlett |
| | NonGaussian (Mild) | Levene |
| Homogeneity of covariance matrices | Gaussian | Box *M* |
| Outliers | Gaussian | Grubbs |
| **Correlation** | | |
| One sample | Gaussian | *z*-test after Fisher *z* transformation |
| Comparison of two independent groups | Gaussian | *z*-test after Fisher *z* transformation |
| In repeated measures | Gaussian | Mauchly |
| Autocorrelation | Gaussian | Durbin–Watson |
| **Survival Curve** | | |
| Comparison of two independent groups | Nonparametric – Large samples | |
| | Same weight to all time points | Log-rank |
| | Weight proportional to $n_t$ | Breslow |
| | Weight proportional to $\sqrt{n_t}$ | Tarone–Ware |
| **Distribution** | | |
| One sample | Nonparametric – Large sample | Kolmogorov–Smirnov |
| | Gaussian or nonGaussian – Small sample | Anderson–Darling |
| | Gaussian – Moderate sample | Shapiro–Wilk |
| Two samples | Nonparametric – Large samples | Kolmogorov–Smirnov |