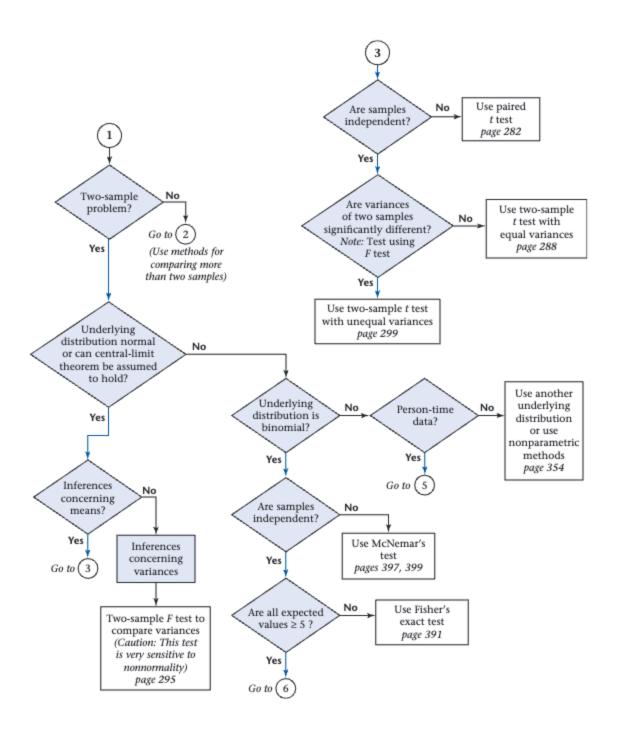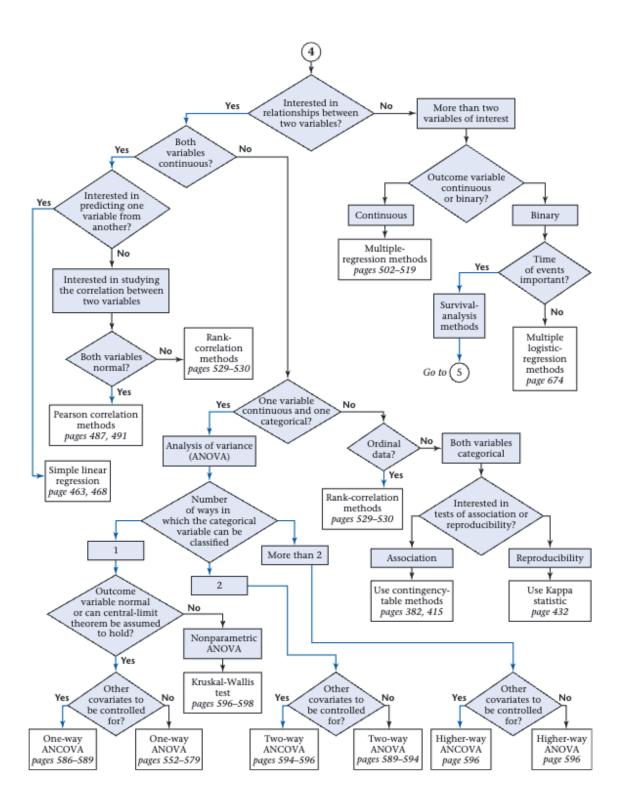# METHODS OF STATISTICAL INFERENCE

```
                              ┌─────────┐
                              │  Start  │
                              └─────────┘
                                   │
                                   ▼
                              ╱─────────╲
                   Yes       ╱ Only one  ╲      No      Go to
              ◄─────────────◄  variable   ►─────────────► ④
              │              ╲ of interest?╱
              │               ╲─────────╱
              ▼
         ╱─────────╲                       Go to
        ╱ One-sample ╲        No            ①
       ◄  problem?    ►──────────────────►
        ╲           ╱
         ╲─────────╱
              │ Yes
              ▼
         ╱──────────────╲
        ╱   Underlying    ╲
       ╱ distribution normal╲    No
  Yes ◄  or can central-limit ►─────────────┐
      │ ╲ theorem be assumed ╱              │
      │  ╲     to hold?     ╱               │
      │   ╲──────────────╱                  ▼
      │                            ╱──────────────╲         ╱──────────────╲      ┌──────────────┐
      ▼                           ╱  Underlying    ╲   No   ╱  Underlying    ╲  No │ Use another  │
 ╱──────────╲     No             ◄ distribution is  ►──────◄ distribution is  ►───►│ underlying   │
╱ Inference  ╲──────┐            ╲  binomial?      ╱        ╲  Poisson?      ╱      │ distribution │
╲ concerning ?╱     │             ╲──────────────╱          ╲──────────────╱       │ or use       │
 ╲──────────╱       ▼                   │ Yes                    │ Yes            │ nonparametric│
  │ Yes      ┌────────────┐             ▼                        ▼                │ methods      │
  │          │ Inference  │      ┌────────────┐          ┌────────────┐           │ pages 341,348│
  │          │ concerning │      │ One-sample │          │ One-sample │           └──────────────┘
  │          └────────────┘      │ binomial   │          │ Poisson    │
  │                │             │ test       │          │ test       │
  │                ▼             └────────────┘          │ pages 261, │
  │          ┌────────────┐            │                 │ 263, 264   │
  │          │ One-sample │            │                 └────────────┘
  │          │ chi-square │            │
  │          │ test for   │            │
  │          │ variances  │            ▼
  │          │(Caution:   │      ╱──────────────╲      ┌──────────┐
  │          │ This test  │     ╱   Normal        ╲ No │ Exact    │
  │          │ is very    │    ◄ approximation     ►──►│ methods  │
  │          │ sensitive  │     ╲   valid?        ╱    │ page 253 │
  │          │ to non-    │      ╲──────────────╱     └──────────┘
  │          │ normality) │            │ Yes
  │          │ pages      │            ▼
  │          │ 246-247    │      ┌────────────┐
  │          └────────────┘      │Normal-theory│
  ▼                              │ methods    │
╱──────────╲     No              │ pages      │
╱ known?    ╲──────┐             │ 250-251    │
╲          ╱       │             └────────────┘
 ╲────────╱        │
  │ Yes            ▼
  ▼            ┌────────────┐
┌────────────┐ │ One-sample │
│ One-sample │ │   t test   │
│   z test   │ │ pages 224, │
│ pages      │ │   225      │
│ 228-229    │ └────────────┘
└────────────┘
```

③

Are samples independent? **No** → Use paired t test *page 282*

Yes ↓

Are variances of two samples significantly different? *Note: Test using F test* **No** → Use two-sample t test with equal variances *page 288*

Yes ↓

Use two-sample t test with unequal variances *page 299*

①

Two-sample problem? **No** → *Go to* ② *(Use methods for comparing more than two samples)*

Yes ↓

Underlying distribution normal or can central-limit theorem be assumed to hold? **No** →

Yes ↓

Inferences concerning means? **No** → Inferences concerning variances → Two-sample F test to compare variances *(Caution: This test is very sensitive to nonnormality)* *page 295*

Yes ↓

Go to ③

Underlying distribution is binomial? **No** → Person-time data? **No** → Use another underlying distribution or use nonparametric methods *page 354*

Yes ↓   **Yes** ↓

Go to ⑤

Are samples independent? **No** → Use McNemar's test *pages 397, 399*

Yes ↓

Are all expected values ≥ 5 ? **No** → Use Fisher's exact test *page 391*

Yes ↓

Go to ⑥

4

Interested in relationships between two variables?

— Yes → Both variables continuous?
— No → More than two variables of interest

More than two variables of interest

Outcome variable continuous or binary?
— Continuous → Multiple-regression methods *pages 502–519*
— Binary → Time of events important?
  — Yes → Survival-analysis methods → *Go to* 5
  — No → Multiple logistic-regression methods *page 674*

Both variables continuous?
— Yes → Interested in predicting one variable from another?
— No → One variable continuous and one categorical?

Interested in predicting one variable from another?
— Yes → Simple linear regression *page 463, 468*
— No → Interested in studying the correlation between two variables

Interested in studying the correlation between two variables

Both variables normal?
— No → Rank-correlation methods *pages 529–530*
— Yes → Pearson correlation methods *pages 487, 491*

One variable continuous and one categorical?
— Yes → Analysis of variance (ANOVA)
— No → Ordinal data?
  — Yes → Rank-correlation methods *pages 529–530*
  — No → Both variables categorical

Both variables categorical

Interested in tests of association or reproducibility?
— Association → Use contingency-table methods *pages 382, 415*
— Reproducibility → Use Kappa statistic *page 432*

Analysis of variance (ANOVA)

Number of ways in which the categorical variable can be classified
— 1
— 2
— More than 2

1 → Outcome variable normal or can central-limit theorem be assumed to hold?
— No → Nonparametric ANOVA → Kruskal-Wallis test *pages 596–598*
— Yes → Other covariates to be controlled for?
  — Yes → One-way ANCOVA *pages 586–589*
  — No → One-way ANOVA *pages 552–579*

2 → Other covariates to be controlled for?
— Yes → Two-way ANCOVA *pages 594–596*
— No → Two-way ANOVA *pages 589–594*

More than 2 → Other covariates to be controlled for?
— Yes → Higher-way ANCOVA *page 596*
— No → Higher-way ANOVA *page 596*

(5)

One-sample problem? —Yes→ Use one-sample test for incidence rates *pages 779–780*

↓ No

Incidence rates remain constant over time?

→ Yes → **Two-sample problem?**
- Yes → Use two-sample test for comparison of incidence rates, if no confounding is present; or methods for stratified person-time data, if confounding is present *pages 783, 786, 794*
- No → Interested in test of trend over more than two exposure groups → Use test of trend for incidence rates *page 806*

→ No → Use survival-analysis methods

↓

Interested in comparison of survival curves of two groups with limited control of covariates?
- Yes → Use log-rank test *page 821*
- No ↓

Interested in effects of several risk factors on survival

↓

Willing to assume several curve comes from a weibull distribution?
- Yes → Use parameter survival methods based on weibull distribution *pages 847–854*
- No → Use Cox proportional-hazards model *page 825*

6

2 × 2 contingency table? —Yes→ *Go to* **A**

No ↓

2 × k contingency table? —Yes→ *Go to* **B**

No ↓

$R \times C$ contingency table, $R > 2$, $C > 2$

↓

Use chi-square test for $R \times C$ tables
page 415

2

Underlying distribution normal or can central-limit theorem be assumed to hold? —Yes→ One-way ANOVA *pages 552–579*

No ↓

Categorical data? —Yes→ Use $R \times C$ contingency-table methods *page 415*

No ↓

Use another underlying distribution or use nonparametric methods such as Kruskal-Wallis test *pages 596–598*

A

↓

Use two-sample test for binomial proportions, or 2 × 2 contingency-table methods if no confounding is present, or the Mantel-Haenszel test if confounding is present *pages 374, 382, 384*

B

↓

Interested in trend over k binomial proportions? —Yes→ Use chi-square test for trend, if no confounding is present, or the Mantel Extension test if confounding is present *pages 418, 670*

No ↓

Use chi-square test for heterogeneity for 2 × k tables *page 415*