|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 1: Types of probability | |  | 2.4: Types of Studies | | | | | |
| Marginal probability | or | **Type** | **Description** | **Estimation** | | **Associated Distribution** | **Fixes** |
| Joint probability |  | Cross-sectional | * Take sample and record where person falls. | Joint *Y*|*X* |  | *X* ~ *Mult*(*n*, *π*) | *n* |
| Conditional probability |  | Prospective | * Take sample, wait *t*, and record where person falls. * Clinical trials: random assignment * Cohort study: participants pick condition | *Y*|*X* | *X* ~ *B*(*n*, *π*) | Row sums (*n*1+ and *n*2+) |
| Total probability |  | Retrospective | * Sample those with or without. * Have enough cases for rare conditions. | *X*|*Y* | *X* ~ *B*(*n*, *π*) | Column sums (*n*+1 and *n*+2) |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 1: Laws of probability | | | | | | | | | | |  |  |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | | **Contingency Table** | | | | | |  | **2 × 2 Sampling Designs** | | | |  | **Y** | | | | **∑**  Marginal | **Distr** | **Fix** | **Notes** | | **X** | *i* × *j* | 1 | 2 | J | *X* ~ *Pois*() | *t* | All values are *nij* | | 1 | *n*11 | *n*12 | *n*1j | *n*1+ | Indep. Bernoulli | One margin | Can estimate joint *Y*|*X* but not cond. | | 2 | *n*21 | *n*22 | *n*2j | *n*2+ | *X* ~ *Mult*(*n*, *π*) | *n* | Not totals | | I | *n*i1 | *n*i2 | *n*ij | *n*i+ | Pseudo-indep. Bernoulli | Treat data as independent *X* ~ *B* | | **∑**  Marginal | | *n*+1 | *n*+2 | *n*+j | *n* | Analysis | | | | | | | | | | | |
| Law of total probability | | | |  | | | | | |
| Addition law of probability | | | |  | | | | | |
| Multiplicative law of probability | | | |  | | | | | |
| Independence | | | | *X* and *Y* are independent if | | | | | |
| Bayes’ Theorem | | | |  | | | | | |
| **Distr.** | | **Parameters** | | | **Probability Function** | | | | | | **Statistics** | | | | 2.10.1: Simpson’s Paradox | | | | | |
| r.v. *X* | |  | | |  | | | | | |  | | | | * General idea: sometimes a trend within a group reverses when multiple groups are combined, due to the *n* values for each group. * *p*(*S*| *X* = 1) + *p*(*S*| *X* = 2) ≠  *p*(*S*| *X* = 1, Z = 1) + *p*(*S*| *X* = 2, Z = 1) * Moral   + Do NOT collapse tables (i.e., use marginal tables) unless appropriate.     - If relationship between *X* and *Y* is the same in marginal tables as in partial (*X* and *Y* with constant *Z*) tables.     - Estimated probabilities need to be about the same in all 3 tables.   + Record control variables that may influence the relationship between *Y* and *X*. | | | | | |
| *Y* ~ *Ber*(*π*) | | Bernoulli trials: *Y*  = {0 = *F*, *S* = 1} | | |  | | | | | |  | | | |
| *X* ~ *B*(*n*, *π*) | | *n* identical indep. Bernoulli trials | | |  | | | | | |  | | | |
| *X* ~ *Mult*(*n*, *π*) | | *k* > 2 outcomes *i*  *n* identical, independent mult. trials | | |  | | | | | |  | | | |
| *X* ~ *Pois*() | | Fixed time or spaces *t*  *X* independent trials with outcome *x*  Useful for big *n* small *π* | | |  | | | | | | Overdisp. when | | | | Make table (*k*, *f*, *p*(*X* = *k*), E(*x*) = *kp*(*X* = *k*))  Use ∑E(*x*) as *µ*  Make table (*j*, *p*(*Y* = *j*) [*Y* ~ *Pois*()], E(*f*) = *np*(*Y* = *j*)) | | | | | |
| *X ~ Hyp(N, r, n)* | | *N* dependent trials{*S* = *r*, *F* = *N* – r} | | |  | | | | | |  | | | |  | | | | | |
| MLE’s  *Not a PDF* | | Known *x*; estimated π | | |  | | | | | |
| Prop. Sampling Distribution  *X* ~ *B* (*n*, *π*) | |  | | |  | | | | | |  | | | |
| 2.3 Association in 2 × 2 Tables | | | | | | | | | | | | | | |  | 2.7: Residuals for Cells in a Contingency Table | | | | |
| Independent binomial sampling | | Compare conditional probabilities | | |  | | | | | | | | | | **Statistics** | | | **Conditions** | |
|  | | | | | | | | | |  | | | 1. Cells where *z* > 2 or *z* > 3 do not exhibit independence. | |
|  | | | | | | | Indep: | | |
| If and | | |  | | |
| **Confidence Intervals** | | | | | | | | | | | | | | | | | | | | |
|  | 1.3.3.1: CI for | | | 2.3.1: [Large Sample] CI for | | | | 2.3.2: Confidence Interval for RR | | | | | 2.3.5: Confidence Interval for θ | | | | | | | |
| **Interval** |  | | |  | | | |  | | | | |  | | | | | | | |
| **SE**  **Estimate** |  | | |  | | | |  | | | | |  | | | | | | | |
| **Hypothesis Tests** | | | | | | | | | | | | | | | | | | | | |
| **Conditions** | | | **H0** | | | **H*a*** | | | **RR (or *p*-value)** | | | | | **TS** | | | | | | |
| 1.3.4: [Large Sample] Approximate Test of Hypothesis for | | | | | | | | | | | | | | | | | | | | |
|  | | |  | | |  | | |  | | | | |  | | | | | | |
|  | | |
|  | | |  | | | | |
| 1.3.6: [Small Sample] Exact Test of Hypothesis for π | | | | | | | | | | | | | | | | | | | | |
|  | | | *π* = *π*0 | | |  | | | Problem-dependent | | | | |  | | | | | | |
|  | | |
|  | | |
| 1.4: [Large Sample] Goodness-of-Fit Test | | | | | | | | | | | | | | | | | | | | |
|  | | | Independence | | | At least one . | | |  | | | | |  | | | | | | |
| 2.3.1: [Large Sample] Difference of Proportions Test for | | | | | | | | | | | | | | | | | | | | |
|  | | |  | | |  | | |  | | | | |  | | |  | | | |
|  | | |
|  | | |  | | | | |
| 2.5: Goodness-of-Fit Test | | | | | | | | | | | | | | | | | | | | |
|  | | |  | | | At least one . | | |  | | | | |  | | |  | If not given all , must estimate *t* parameters.   1. Set 2. Solve for *t* unknowns. 3. Make 2 × 2 table of s; find column totals and equate to s. 4. Solve for unknown values. | | |
| 2.6: [Large Sample] Tests for Independence (*χ*2 and *G*2) | | | | | | | | | | | | | | | | | | | | |
|  | | | Independence | | | At least one  Dependence | | | *t* = *I* + *J* – 2 | | | | |  | | |  | Do not partition | | |
|  | | | Partition by combining columns  If each is independent. | | |
| 2.8: [Small Sample] Test for Independence | | | | | | | | | | | | | | | | | | | | |
|  | | | Independence | | |  | Dependence | | Generally add values of *p* where | | | | |  | | | is observed | | | 1. Create all possible tables given a value of . 2. Compute TS’s. |
|  |
|  |