• PROBABILITY

• $P(A\cap B) = P(A|B)\cdot P(B)$

↳ if A, B independent: $P(A\cap B) = P(A)\cdot P(B)$

• $P(A|B) = \frac{P(A\cap B)}{P(B)} = \frac{P(A)\cdot P(B|A)}{P(B)}$

• $P(A\cup B) = P(A) + P(B) - P(A\cap B)$

↳ if A, B mutually exclusive: $P(A\cup B) = P(A) + P(B)$

• $(A\cup B)' = A'\cap B'$

• $P(A\cup B\cup C) = P(A) + P(B) + P(C) - P(A\cap B) - P(A\cap C) - P(B\cap C) + P(A\cap B\cap C)$

• $P(E_1\cup\ldots\cup E_m) = \sum_{i=1}^{m} P(E_i) - \sum_{i<j} P(E_i\cap E_j) + \sum_{i<j<k} P(E_i\cap E_j\cap E_k) + \ldots + (-1)^{m} P(E_1\cap\ldots\cap E_m)$

• COMBINATION (unordered subset): $\binom{m}{k} = \frac{m!}{k!\,(m-k)!}$ — size of set — NO R; $\binom{m+k-1}{k}$ — WITH R

• PERMUTATION (ordered subset): $\frac{m!}{(m-k)!}$ — size of subset — NO R; m^k — WITH R

BAYES RULE:

• $P(A_k|B) = \frac{P(A_k\cap B)}{P(B)} = \frac{P(B|A_k)\cdot P(A_k)}{\sum P(B|A_i)\cdot P(A_i)}$

• DISCRETE RANDOM VARIABLES

• p.m.f: $p(x) = \mathbb{P}(X=x) \overset{\Sigma}{=} 1$ $\quad \sum p(y)$

• c.d.f: $F(x) = \mathbb{P}(X\le x) \to \mathbb{P}(a\le x\le b) = F(b) - F(a)$

• $E(x) = \mu_x = \sum x\cdot p(x)$

$E[h(x)] = \sum h(x)\cdot p(x)$

$E(aX+b) = aE(x) + b$

$E[g_1(x)+\ldots+g_m(x)] = E[g_1(x)] + \ldots + E[g_m(x)]$

• $V(x) = \sum (x-\mu)^2\cdot p(x) = E[(X-\mu)^2]$

$V(x) = \sigma^2 = E(x^2) - [E(x)]^2$

$V(aX+b) = a^2 V(x) = a^2\sigma_x^2$

• BERNOULLI (experiment with 2 outcomes)

$Y=1$ if success, $Y=0$ if not

$f_X(p) = p^x(1-p)^{1-x}$ $\quad E(x) = p$ $\quad V(x) = p(1-p)$

$P(Y) = \begin{cases} p, & y=1 \\ 1-p, & y=0 \end{cases}$

• BINOMIAL (n° of successes in N trials) $b(x; m, p)$

extraction WITH REPLACEMENT

$p(x) = \binom{m}{x} p^x (1-p)^{m-x}$; $x = 0, 1, \ldots, m$ $\quad E(x) = mp$ $\quad V(x) = mp(1-p)$

↳ can be approximated with a NORMAL distribution $\to Bin(x; m, p) = \Phi\left(\frac{x+0{,}5-mp}{\sqrt{mp(1-p)}}\right)$

• HYPERGEOMETRIC

extraction WITHOUT REPLACEMENT

$h(x; m, M, N)$

N: size of population

M: # of success (N−M: # FAILURE)

m: size of the sample

$\mathbb{P}(X=x) = \frac{\binom{M}{x}\cdot\binom{N-M}{m-x}}{\binom{N}{m}}$ $\quad E(x) = m\cdot\frac{M}{N}$ $\quad V(x) = \frac{N-m}{N-1}\cdot m\cdot\frac{M}{N}\cdot\left(1-\frac{M}{N}\right)$

↳ can be approximated with a NEGATIVE BINOMIAL distribution

• NEGATIVE BINOMIAL (n° of failures before the r-th success)

$mb(x; x, p) = (1-p)^x\cdot p^r\cdot\binom{x+r-1}{r-1}$ $\quad$ r: # of success; p: $\mathbb{P}$(success)

$E(x) = \frac{r(1-p)}{p}$ $\quad V(x) = \frac{r(1-p)}{p^2}$

• DISCRETE UNIFORM DISTRIBUTION

$f(x) = f(x; N) = \begin{cases} \frac{1}{N} & x = 1, 2, \ldots, N \\ 0 & \text{otherwise} \end{cases}$ $\quad E(x) = \frac{N+1}{2}$ $\quad V(x) = \frac{N^2-1}{12}$

• POISSON

binomial with $m > 50$ e $p < 5$ $\longrightarrow p(x;\mu)$, $\mu = mp$

$Poi(x;\mu) = \frac{e^{-\mu}\mu^x}{x!}$ $\quad x = 0, 1, 2, \ldots$ $\quad E(x) = V(x) = \mu$

~ Poisson process: occurrance of events OVER TIME; $\quad \alpha$: rate of the process $\sim Poi(\alpha t)$

• STATISTICAL METHODS

$X_m \xrightarrow{a.s.} X \Rightarrow X_m \xrightarrow{P} X \Rightarrow X_m \xrightarrow{d} X$ (LLN: $\xrightarrow{P}$; CLT: $\xrightarrow{d}$)

$\Updownarrow \qquad \Updownarrow$

$X_m - X \xrightarrow{a.s.} \emptyset \Rightarrow X_m - X \xrightarrow{P} \emptyset$

$\Updownarrow$

$X_m - X \xrightarrow{d} \emptyset$

• Chebyshev Inequality

$\mathbb{P}[|X|\ge a] \le \frac{E[|X|^K]}{a^K}$ $\quad \forall a>0$, $\forall K\in\mathbb{N}^+$

$\mathbb{P}[|X|< a] \ge 1 - \frac{E[|X|^K]}{a^K}$

• Corollary

$\mathbb{P}[|X-\mu|<a] \ge 1 - \frac{Var[X]}{a^2}$

• CONTINUOUS RANDOM VARIABLES

• p.d.f.: $P(a\le x\le b) = \int_a^b f(x)\,dx \longrightarrow P(a\le x\le b) = P(a<x<b) = P(a\le x<b) = P(a<x\le b)$

• c.d.f.: $F(x) = P(X\le x) = \int_{-\infty}^{x} f(y)\,dy$ $\quad$ • $\mu_x = E(x) = \int_{-\infty}^{+\infty} x\cdot f(x)\,dx \to E(ax+b) = a\mu_x + b$

• $\sigma^2 = V(x) = \int_{-\infty}^{+\infty} (x-\mu)^2 f(x)\,dx = E[(X-\mu)^2]$ $\quad V(x) = E(x^2) - [E(x)]^2$ $\quad V(ax+b) = a^2\sigma_x^2$

• UNIFORM DISTRIBUTION

$E(x) = \frac{1}{2}(a+b)$ $\quad V(x) = \frac{1}{12}(b-a)^2$ $\quad f(x; A, B) = \begin{cases} \frac{1}{B-A}, & A\le x\le B \\ \emptyset, & \text{otherwise} \end{cases}$

• NORMAL DISTRIBUTION

STANDARD $\to Z = f(z; 0, 1)$

$f(x; \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}}\cdot\exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$ $\quad E(x) = \mu$ $\quad V(x) = \sigma^2$

• MULTINOMIAL DISTRIBUTION

p_i = outcome "i" in any particular trial

X_i = # of trials with outcome "i"

$P(X_1, X_2, \ldots, X_m) = \begin{cases} \frac{m!}{(X_1!)(X_2!)\ldots(X_m!)}\, p_1^{X_1}\cdot p_2^{X_2}\cdots p_m^{X_m} & (m = X_1+X_2+\ldots+X_m) \\ 0 & \text{otherwise} \end{cases}$

e.i.: $p_1 = p_3 = 0.25$, $p_2 = 0.5$

$P(X_1=2, X_2=5, X_3=3) = P(2,5,3) = \frac{10!}{2!\,5!\,3!}\cdot 0{,}25^2\cdot 0{,}5^5\cdot 0{,}25^3 = 0{,}0769$

• EXPONENTIAL DISTRIBUTION

$f(x;\lambda) = \begin{cases} \lambda e^{-\lambda x}, & x\ge 0 \\ 0, & \text{otherwise} \end{cases}$ $\quad F(x) = (x;\lambda) = \begin{cases} 0, & x\le 0 \\ 1-e^{-\lambda x}, & x>0 \end{cases}$ $\quad \mu = \frac{1}{\lambda}$ $\quad \sigma^2 = \frac{1}{\lambda^2}$

• GEOMETRIC DISTRIBUTION

$f(x; K, p) = p(1-p)^{K-1}$ $\quad E(x) = \frac{1}{p}$ $\quad V(x) = \frac{1-p}{p^2}$

• BETA DISTRIBUTION (generalization of uniform distribution)

$Beta(\alpha, \beta)$

$f_p(t) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\cdot\Gamma(\beta)}\cdot p^{\alpha-1}\cdot(1-p)^{\beta-1}$ $\quad p\in[0,1]$ $\quad$ (GAMMA)

• if $\alpha, \beta = 1 \to f_p(t) = 1$, $p\in[0,1] \to X\sim U(0,1)$

$E[Beta(\alpha,\beta)] = \frac{\alpha}{\alpha+\beta}$ $\quad V[Beta(\alpha,\beta)] = \frac{\alpha\cdot\beta}{(\alpha+\beta)^2\cdot(\alpha+\beta+1)}$ $\quad \Gamma(\alpha) = (\alpha-1)!$

if we don't have knowledge of the problem, as the prior $\longrightarrow \pi\sim Beta(1,1) \longrightarrow$ Uniform

• JOINT PROBABILITY DISTRIBUTION

• Joint p.m.f.: $p(x,y) = P(X=x \text{ AND } Y=y)$

• $P[(X,Y)\in A] = \sum\sum_{(x,y)\in A} P(x,y)$

• $P[(X,Y)\in A] = \iint_A f(x,y)\,dx\,dy$

$\to$ X, Y INDEPENDENT $\to$

$E[g(x)h(y)] = E[g(x)]\cdot E[h(y)]$

$p(x,y) = p_x(x)\cdot p_y(y)$

$f(x,y) = f_x(x)\cdot f_y(y)$

• COVARIANCE

$Cov(X,Y) = E(X\cdot Y) - \mu_x\cdot\mu_y \to$ INDEP $\to Cov(X,Y) = \emptyset$

$Cov(X,X) = V(x)$

$Cov(aX+b, Y) = a\cdot Cov(X,Y)$

Conditional: $f_{Y|X}(y|x) = \frac{f(x,y)}{f_X(x)}$

• CORRELATION

$\rho_{XY} = \frac{Cov(X,Y)}{\sigma_X\cdot\sigma_Y}$

$Corr(aX+b, cY+d) = Corr(X,Y)$

$-1\le Corr(X,Y)\le 1$

if X, Y INDEP. $\to Corr(X,Y) = \emptyset$ [does NOT imply independence]

• LAW OF LARGE NUMBERS (LLN):

X_i INDEPENDENT and IDENTICALLY DISTRIBUTED, defined: $\bar{X}_m = \frac{\sum_i^m X_i}{m} \to$ one has:

$\bar{X}_m - \mu \xrightarrow{P} \emptyset \iff \bar{X}_m \xrightarrow{P} \mu$

• CENTRAL LIMIT THEOREM (CLT):

X_i INDEPENDENT and IDENTICALLY DISTRIBUTED, defined: $S_m = X_1 + \ldots + X_m \to$ one has:

$U_m = \frac{S_n - m\mu}{\sqrt{m\sigma^2}} \xrightarrow{d} Z\sim N(0,1) \to E(S_n) = m\cdot\mu$, $V(S_m) = m\cdot\sigma^2$

↳ IMMEDIATE CONSEQUENCE: for large values of m $(m\ge 30)$

$\bar{X}_m = \frac{S_m}{m} \longrightarrow \bar{X}_m\sim N\left(\mu, \frac{\sigma^2}{m}\right)$

• POINT ESTIMATION

$\cdot\ \hat{\mu} = \frac{1}{m}\sum_{i}^{m} X_i = \bar{X}_m \qquad \cdot\ \hat{\sigma}^2 = \frac{\sum (X_i - \bar{X}_m)^2}{m-1}$

* MOMENTS: $M_1 = \frac{1}{m}\sum_{i=1}^{m} X_i \xrightarrow{P} E[X]$

$M_2 = \frac{1}{m}\sum_{i=1}^{m} X_i^2 \xrightarrow{P} E[X^2]$

$\vdots$

$\cdot\ X \sim Exp(\lambda) \rightarrow \hat{\lambda} = \frac{1}{M_1} = \frac{m}{\sum X_i}$

$\cdot\ X \sim Bin(K, p) \rightarrow \hat{K} = \frac{M_1^2}{M_1 + M_1^2 - M_2}$

$\hat{p} = \frac{M_1}{\hat{K}}$

* LIKELIHOOD FUNCTION: $L(\theta | X_1, \dots, X_m) = \prod_{i=1}^{m} f_\theta(X_i)$ [ℓn, derivata = 0]

$\hat{\Theta} = \text{argmax}\{L(\theta | X_1, \dots, X_m)\}$

$\cdot\ X \sim Bern(p) \rightarrow \hat{p} = \frac{\sum X_i}{m} = \bar{X}_m$

* BAECYSIAN ESTIMATION: B – geom; Gamma – Pois; Unif – Bern;

$f_x(x) = \int f_x(x|\theta)\cdot \pi(\theta)\, d\theta = \sum f_x(x|\theta_i)\cdot \pi(\theta_i)$ ($\in \mathbb{R}$)

$\pi(\theta | x) = \frac{f_x(x|\theta)\cdot \pi(\theta)}{f_x(x)}$ (POSTERIOR, LIKELIHOOD)

$\pi(\theta|x) \propto f_x(x|\theta)\cdot \pi(\theta)$ (PRIOR)

• ESTIMATE OF μ:
if $X \sim N(\mu, \sigma^2)$, σ^2 known
if prior $\sim N(\nu, \tau^2)$ $\nu = \frac{\mu_a + \mu_b}{2}$

$\hat{\mu} = \frac{m \cdot \tau^2}{m\tau^2 + \sigma^2}\bar{X}_m + \frac{\sigma^2}{m\tau^2 + \sigma^2}\nu$

• BIAS: $Bias(\hat{\Theta}) = E[\hat{\Theta}] - \theta$

if $E[\hat{\Theta}] = \theta \longrightarrow \hat{\Theta}$ is UNBIASED

• EFFICIENCY: $V[\hat{\Theta}_1] \le V[\hat{\Theta}_2]$ if $\hat{\Theta}_1$ more efficient than $\hat{\Theta}_2$

• MEAN SQUARE ERROR: $MSE(\hat{\Theta}) = E[(\hat{\Theta} - \theta)^2] = V[\hat{\Theta}] + Bias(\hat{\Theta})^2$

• CRAMER-RAO LOWER BOUND: let X have density $f_x(t, \theta)$ with continuous I AND II derivative, then for unbiased estimator $\hat{\Theta}$:

$$V[\hat{\Theta}] \ge \left\{ m \cdot E\left[\left(\frac{\partial \ell n f_x(t,\theta)}{\partial \theta}\right)^2\right]\right\}^{-1} = \left\{ m \cdot E\left[\left(\frac{\partial^2 \ell n f_x(t,\theta)}{\partial \theta}\right)\right]\right\}^{-1}$$

• CONSISTENCY: $\hat{\Theta} = W(X_1, \dots, X_m)$ is CONSISTENT for θ if:

$\lim_{m\to\infty} P[|\hat{\Theta}_m - \theta| < \varepsilon] = 1$

$f_{Y_{MAX}}(y) = m \cdot F_Y(y)^{m-1} \cdot f_Y(y) \longrightarrow F_Y(y) = \int_0^y f_Y(t)\cdot dt$

$f_{Y_{MIN}}(y) = m \cdot f_Y(y) \cdot (1 - F_Y(y))^{m-1}$

• INTERVAL ESTIMATION

• $100(1-\alpha)\%$ CI

μ

1) mo N, Known σ^2, Large sample

$I = [\bar{X}_m - Z_{1-\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{m}} ; \bar{X}_m + Z_{1-\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{m}}]$

2) mo N, NOT Known σ^2, Large sample

$I = [\bar{X}_m - Z_{1-\frac{\alpha}{2}} \cdot \frac{\hat{\sigma}}{\sqrt{m}} ; \bar{X}_m + Z_{1-\frac{\alpha}{2}} \cdot \frac{\hat{\sigma}}{\sqrt{m}}]$

3) N, Known σ^2, Small sample

$I = [\bar{X}_m - Z_{1-\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{m}} ; \bar{X}_m + Z_{1-\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{m}}]$

4) N, NOT Known σ^2, Small sample

$I = [\bar{X}_m - t_{1-\frac{\alpha}{2}} \cdot \frac{\hat{\sigma}}{\sqrt{m}} ; \bar{X}_m + t_{1-\frac{\alpha}{2}} \cdot \frac{\hat{\sigma}}{\sqrt{m}}]$

t-student with $(m-1)$ degrees of freedom

σ^2

1) only N

$I = \left[\frac{(m-1)\hat{\sigma}_m^2}{q_{1-\frac{\alpha}{2}}} ; \frac{(m-1)\hat{\sigma}_m^2}{q_{\frac{\alpha}{2}}}\right]$

q – chi square with $m-1$ degrees of freedom

d

$\bar{X}_m - \bar{Y}_m \sim N\left(\mu_x - \mu_y ; \frac{\sigma_x^2}{m_x} + \frac{\sigma_y^2}{m_y}\right)$

$I = \left[\bar{X}_m - \bar{Y}_m \pm Z_{\left(1-\frac{\alpha}{2}\right)}\sqrt{\frac{\sigma_x^2}{m_x} + \frac{\sigma_y^2}{m_y}}\right]$

we can use $\hat{\sigma}_x$ e $\hat{\sigma}_y$ instead of σ_x, σ_y

• EXPONENTIAL CI

let $X_i \sim Exp(\lambda)$ INDEP.

$Y = \sum X_i \sim \Gamma(m, \lambda)$

$\Rightarrow 2\lambda Y \sim \chi^2_{2m}$

$I_{1-\alpha} = \left[\frac{\chi^2_{2m, \frac{\alpha}{2}}}{2\sum X_i} , \frac{\chi^2_{2m, 1-\frac{\alpha}{2}}}{2\sum X_i}\right]$

• BERNOULLI CI

let $X \sim Bern(p)$ $Y \sim Bin(p, m)$

$\Rightarrow \frac{Y - mp}{\sqrt{mp\cdot(1-p)}} \sim N(0,1)$

$I_{1-\alpha} = \left[\hat{p} - Z_{1-\frac{\alpha}{2}}\sqrt{\frac{\hat{p}(1-\hat{p})}{m}} , \hat{p} + Z_{1-\frac{\alpha}{2}}\sqrt{\frac{\hat{p}(1-\hat{p})}{m}}\right]$

• DELTA METHOD

1. Estima $\hat{\alpha}$ or → riscrivi l'estimator come generica funzione in X
2. I° derivata di $g(x) \longrightarrow g(\mu)$ e $g'(\mu)$
3. $g(\bar{X}_m) \xrightarrow{d} N\left(g(\bar{X}_m), g'(\bar{X}_m)^2 \cdot \frac{\sigma^2}{m}\right)$

• INTERV. ESTIM. BAYESIAN FRAMEWORK

$\pi(\theta \in [L(\bar{x}), U(\bar{x})] | \bar{x}) = 1 - \alpha$

$I_{1-\alpha} = \left[E[\mu|\bar{x}] - Z_{1-\frac{\alpha}{2}}\sqrt{V[\mu|\bar{x}]}, E[\mu|\bar{x}] + Z_{1-\frac{\alpha}{2}}\sqrt{V[\mu|\bar{x}]}\right]$

• HYPOTHESIS TESTING

α: type I Error: rejecting H_0 when it is true
β: type II Error: not rejecting H_0 when it is false
] decreasing α, β increases; if m increases, α, β decrease

• Confidence level α: $\alpha = P(\text{type I error})$ $\qquad \alpha \equiv$ p-value (p.a sim, $\geq$)

• Rejection region: $c = (-\infty, \frac{\alpha}{2}) \cup (1 - \frac{\alpha}{2}, +\infty)$

μ $\quad H_0: \mu = \mu_0 \quad H_1: \mu \neq \mu_0$

1) mo N, Known σ^2, Large sample

$Z_m = \frac{\bar{X}_m - \mu_0}{\sigma / \sqrt{m}} \quad c = (-\infty, -Z_{1-\frac{\alpha}{2}}) \cup (Z_{1-\frac{\alpha}{2}}, \infty)$

2) mo N, NOT Known σ^2, Large sample

$Z_m = \frac{\bar{X}_m - \mu_0}{\hat{\sigma} / \sqrt{m}} \quad c = (-\infty, -Z_{1-\frac{\alpha}{2}}) \cup (Z_{1-\frac{\alpha}{2}}, \infty)$

3) N, Known σ^2, Small sample ✶

$Z_m = \frac{\bar{X}_m - \mu_0}{\sigma / \sqrt{m}} \quad c = (-\infty, -Z_{1-\frac{\alpha}{2}}) \cup (Z_{1-\frac{\alpha}{2}}, \infty)$

4) N, NOT Known σ^2, Small sample

$T_m = \frac{\bar{X}_m - \mu_0}{S_n / \sqrt{m}} \quad c = (-\infty, -t_{1-\frac{\alpha}{2}}) \cup (t_{1-\frac{\alpha}{2}}, \infty)$

t-student with $(m-1)$ degrees of freedom

σ^2 $\quad H_0: \sigma^2 = \sigma_0^2 \quad H_1: \sigma^2 \neq \sigma_0^2$

1) only N

$Q_m = \frac{(m-1)\cdot S_m^2}{\sigma_0^2} \quad c = [0, q_{\frac{\alpha}{2}}) \cup (q_{1-\frac{\alpha}{2}}, +\infty)$

q – chi square with $m-1$ degrees of freedom

d $\quad H_0: \mu_x = \mu_y \quad H_1: \mu_x \neq \mu_y$

$\mu = \mu_x - \mu_y \rightarrow H_0: \mu = 0$

1) $\tilde{T} = \frac{(\bar{X}_m - \bar{Y}_m) - \mu}{\sqrt{\frac{(m+m)[(m-1)\cdot\hat{\sigma}_{x,m}^2 + (m-1)\hat{\sigma}_{y,m}^2]}{m\cdot m\cdot(m+m-2)}}}$

$t_{m+m-2} \rightsquigarrow Z_{m+m-2}$ (DOF) for large sample

$c = (-\infty, -t_{1-\frac{\alpha}{2}}) \cup (t_{1-\frac{\alpha}{2}}, \infty)$

* LIKELIHOOD RATIO TEST: likelihood function $L(\theta|\bar{X}) = \prod f(x_i|\theta) = \ell n(\theta)$

$H_0: \theta \in \Theta_0 \qquad \lambda(\bar{x}) = \frac{\sup_{\theta_0} L(\theta_0|\bar{X})}{\sup_{\theta} L(\hat{\theta}|\bar{X})} \in [0, 1]$

rejected if $\lambda(\bar{x}) \le \lambda^*$ . . . . for some $\lambda^* \in [0, 1]$

$\alpha = P[\lambda(\bar{X}_m) \le \lambda^* \mid H_0 \text{ is true}]$

* BAYESIAN TEST $\qquad \frac{P(\theta \in \Theta_0 | \bar{X})}{P(\theta \in \Theta_1 | \bar{X})} < c$

* GOODNESS OF FIT (test the shape of the distribution)

$H_0: X \sim f = f_0 \qquad H_1: X \sim f \neq f_0$ $\qquad$ used for both discrete and continuos

• LARGE SAMPLES ARE REQUIRED $(m > 100)$

Let $I_1, \dots, I_N$ be the possible outcomes $\qquad f_i$ = m° times I_i occurs in $\bar{X}$

$p_i = P[X \in I_i \mid f_0 \text{ is true density}]$

then $W = \sum_{i=1}^{N} \frac{(f_i - mp_i)^2}{mp_i}$

has approximatly a χ^2 distribution with $N-1$ dof $(mp_i \ge 5\ \forall i)$

• if parameters are estimated through the sample, W has a χ^2 distribution with $N - m - 1$ dof

• if two RV instead of one: $W = \sum_i \sum_j \frac{(f_{ij} - p_i p_j m)^2}{p_i p_j m}$

INDEPENDENCE TEST $\quad H_0$: IND $\quad H_1$: not IND $\qquad$ degrees of freedom $(row-1)(column-1)$

1. table of probabilities with observed values
2. calculate total for each row, column and overall → N
3. fill in a second table with the expected values $E = \frac{RAW_{TOT} \cdot COL_{TOT}}{N}$
4. calculate total for each row, column and overall for new table
5. calculate $\forall$ cell: $\chi^2_{...} = \frac{\sum (OBS_i - E_i)^2}{E_i}$
6. if ⑤ ≤ critical value ⟶ ⟶ we fail to reject

✶ when approximating $bin \sim N(pm, \sqrt{pqm})$, don't divide by $\sqrt{m}$