- $E[X] = \mu_x = \sum_{x \in D} x \cdot p_X(x)$

  IF $h(X) = aX + b \rightarrow E[h(X)] = a\,E[X] + b$

- $V[X] = \sigma_x^2 = \sum_{D} (x-\mu)^2 \cdot p(x) = E[(x-\mu)^2]$

  $= E[X^2] - E^2[X]$

  $h(X) = aX + b \rightarrow V[h(X)] = a^2\,V[X]$

  $V[X+Y] = V[X] + V[Y] + 2\,Cov(X,Y)$

- $V[aX + bY] = a^2\,V[X] + b^2\,V[Y] + 2ab\,Cov(X,Y)$

▼ **Disposizioni:**

Semplici : $D_{n,k} = \frac{n!}{(n-k)!}$

Ripetzione : $D'_{n,k} = n^k$

▼ **Permutazioni:**

Semplici : $P_n = n!$

Ripetzione : $P_n^{(\ldots,\ldots)} = \frac{n!}{\ldots!\ldots!}$

▼ **Combinazioni:**

Semplici : $C_{n,k} = \binom{n}{k} = \frac{n!}{k!\,(n-k)!}$

Ripetzione : $C'_{n,k} = \binom{n+k-1}{k} = \frac{(n+k-1)!}{k!\,(n-1)!}$

- CHEBYSEV'S INEQUALITY.

  $P(|X-\mu| \geq k\sigma) \leq \frac{1}{k^2}$ | $P(|X-\mu| \leq k) \geq 1 - \frac{V^{\alpha}[X]}{k^{\alpha}}$

- BERNOULLI DISTRIBUTION:

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

  - $E[Y] = p$
  - $V[Y] = p(1-p)$

- BINOMIAL DISTRIBUTION: n TRIALS / EACH TRIAL $X \sim$ BERNOULLI (p) (WITH REPLACEMENT)

  - $X \sim Bin(n, p)$
    $P(X = k) = p(k) = \binom{n}{k} p^k (1-p)^{n-k}$
  - $E[X] = np$
  - $V[X] = np(1-p)$

- HYPERGEOMETRIC DISTRIBUTION: ~> GIVEN POPULATION OF N ELEMENT, I KNOW M SUCCESSES
  ↳ SELECTED $n < N$ ELEMENTS : $X = k$ SUCCESSES ?

  (NO REPLACEMENT)

  - $X \sim$ HYPERGEOMETRIC (n, M, N) :
    $P(X = k) = \frac{\binom{M}{k}\binom{N-M}{n-k}}{\binom{N}{n}}$
  - $E[X] = n \cdot \frac{M}{N}$ / $\frac{M}{N} = p$
  - $V[X] = \left(\frac{N-n}{N-1}\right) \cdot n \cdot \frac{M}{N}\left(1 - \frac{M}{N}\right)$

  { IF $\frac{n}{N} < 5\%$ $\rightarrow h() \approx Bin()$

- NEGATIVE BINOMIAL DISTRIBUTION: ~> EXPERIMENT CONTINUES UNTIL A TOTAL OF r SUCCESSES HAVE BEEN OBSERVED

  X : n° F THAT PRECEDES r^{TH} S, $X \sim nb(r, p)$

  $P(X = k) = \binom{r-1+k}{r-1} p^r (1-p)^k$

  - $E[X] = \frac{r(1-p)}{p}$
  - $V[X] = \frac{r(1-p)}{p^2}$

- IF $r = 1$ $\longrightarrow$ GEOMETRIC DISTRIBUTION: $P(X = k) = p(1-p)^k$
  ↳ AT LEAST 1 SUCCESS

- SUMMARY:

  - $X \sim Bin(n, p)$, $P(X = k)$. k SUCCESSES (S WITH p) IN n TRIALS
  - $X \sim h(n, M, N)$, $P(X = k)$: k SUCCESSES IN SUBSET OF LENGTH n IN A POPULATION OF N ELEMENTS / M SUCCESSES
    ↳ IF $\frac{n}{N} < 5\% \rightarrow h() \approx Bin()$
  - $X \sim nb(r, p)$, $P(X = k)$: k FAILURES BEFORE r^{TH} SUCCESS
    $\xrightarrow{r=1}$ $X \sim$ GEOMETRIC (p), $P(X = k)$: k FAILURE BEFORE 1 SUCCESS

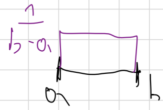
• DISCRETE RV :

- UNIFORM :

$$f_X(x;N) = \begin{cases} 1/N & , x = 1,2,\dots \\ 0 & , \text{OTHERWISE} \end{cases}$$

- $E[X] = \dfrac{N+1}{2}$
- $V[X] = \dfrac{N^2-1}{12}$

• IF CONTINUOUS



- $E[X] = \frac{1}{2}(b+a)$
- $V[X] = \frac{1}{12}(b-a)^2$

- POISSON DISTRIBUTION :

$X \sim Poiss(\mu) \;/\; \mu = E[X] = V[X]$

$$p(x;\mu) = \frac{\mu^x}{x!} e^{-\mu}, \quad x \in [0;\infty)$$

$$\begin{cases} \underset{\text{BINOMIAL}}{b(x;n,p)} \longmapsto \underset{\text{POISSON}}{p(x;\mu)}, \quad \mu = n \cdot p \\ \text{FOR } (n > 50) \text{ AND } (np < 5) \end{cases}$$

- LET $P_k(t) = p$ THAT k EVENTS IN t TIME

$$\rightarrow P_k(t) = \frac{(\alpha t)^k}{k!} e^{-\alpha t} \;/\; \mu = \alpha t$$

- $Y = X_1 + \dots + X_n \;/\; X_i \sim Poiss(\mu)$

$\rightarrow Y \sim Poiss(n\mu)$

• CONTINUOUS RV :

- MEDIAN : $\tilde{\mu} \;/\; F(\tilde{\mu}) = 50\%$

- THE NORMAL DISTRIBUTION :

$$f(x;\mu,\sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}}$$

- $E(X) = \mu$
- $V(X) = \sigma$

$$\begin{cases} \cdot\, N \mapsto Z : \; Z = \dfrac{X-\mu}{\sigma} \\ \cdot\, Z \longmapsto N : \; X = \sigma \cdot Z + \mu \end{cases}$$

- STANDARD NORMAL : $Z \sim N(0,1)$
- $P(Z \geq z_\alpha) = \alpha \;/\; z_\alpha = 100(1-\alpha)^{th}$
- $(100p)^{th}$ PERCENTILE FOR $N(\mu,\sigma^2) = \mu + [(100p)^{th}$ PERCENTILE FOR $Z]\,\sigma$

$$\begin{cases} \text{IF } np \geq 10 \text{ AND } n(1-p) \geq 10 \rightarrow \text{VALID APPROX.} \\ \rightarrow P(X \leq x) = Bin(x;n,p) \approx \phi\left(\dfrac{x+0.5 - \overset{np}{\mu_x}}{\underset{\sqrt{np(1-p)}}{\sigma_x}}\right) \end{cases}$$

- THE EXPONENTIAL DISTRIBUTION :

$$f(x;\lambda) = \begin{cases} \lambda e^{-\lambda x} & , x \geq 0 \\ 0 & , \text{OTHERWISE} \end{cases}$$

- $E[X] = 1/\lambda$
- $V[X] = 1/\lambda^2$

- CDF : $F(x;\lambda) = \begin{cases} 1 - e^{-\lambda x} & , x \geq 0 \\ 0 & , x < 0 \end{cases}$

$\rightarrow$ MEMORYLESS PROPERTY : $P(X > t + t_0 \mid X > t_0) = P(X > t)$

• JOINT RV.

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k} = CMF(Bin(n,k))$$

• 2 RV X, Y are independent $\iff \forall (x,y)$
- $p(x,y) = p_X(x)\, p_Y(y)$ — DISCRETE
- $f(x,y) = f_X(x)\, f_Y(y)$ — CONTINUOUS

MULTINOMIAL EXPERIMENT

• Joint PMF of $X_1, \dots, X_r$: multinomial distribution / n : # identical trials, $\forall$ trial $\to r$ outcomes possible

$$p(x_1, \dots, x_r) = \begin{cases} \dfrac{n!}{(x_1!) \cdot \ldots \cdot (x_r!)} \cdot p_1^{x_1} \cdots p_r^{x_r}, & x_i = 0, 1, \dots ;\ x_1 + \dots + x_r = n \\ 0, & \text{otherwise} \end{cases}$$

• CONDITIONAL DISTRIBUTIONS :

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

• COVARIANCE :

$$Cov(X,Y) = E[(X-\mu_X)(Y-\mu_Y)]$$

• PROPERTIES :
- $Cov(X,Y) = E(XY) - \mu_X \mu_Y$
- X, Y independent $\to Cov(X,Y) = 0$
- $Cov(X,X) = V(X) = E[(X-\mu_X)^2]$
- $Cov(aX+b, Y) = a \cdot Cov(X,Y)$
- $Cov(aX+bY, cV+dW) = ac\, Cov(X,V) + ad\, Cov(X,W) + bc\, Cov(Y,V) + bd\, Cov(Y,W)$
- $Cov(aX+bY) = a\, Cov(X,Y) + b\, V[Y]$

If X, Y indip. :
- $Cov(X, X+Y) = V(X)$

• CORRELATION :

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

• PROPERTIES :
- If $a \cdot c > 0 \to \rho(aX+b,\ cY+d) = \rho(X,Y)$
- $\forall X, Y : -1 \le \rho_{X,Y} \le 1$
- If X, Y independent $\overset{\not\Leftarrow}{\Longrightarrow} \rho_{X,Y} = 0 \to X, Y$ uncorrelated
- $\rho_{X,Y} = \pm 1 \iff Y = aX + b,\ a \neq 0,\ a, b \in \mathbb{R}$

$\to \rho_{X,Y} = 0 \not\to X, Y$ independent : it only says there is not a linear dependence

$\to$ 2 R.V. can be uncorrelated but highly dependent : strong non-linear relation

$\to$ ASSOCIATION $\neq$ CAUSATION :

If $\rho_{X,Y} \gg 0 \to \uparrow X \not\to \uparrow Y$

PART 2 :

- CONVERGENCE IN DISTRIBUTION :

$$\begin{cases} X_n \xrightarrow{d} X \\ F_n \xrightarrow{d} F \end{cases} \text{ OR , FOR } X \sim F \text{ IF } \lim_{n \to \infty} F_n(t) = F(t), \forall t \,/\, F \text{ IS CONTINUOUS IN } t$$

$\rightsquigarrow$ DISTRIBUTIONS

- $X_n \xrightarrow{d} X \iff E[g(X_n)] \xrightarrow{n\to\infty} E[g(X)], \forall g$ BOUNDED AND CONTINUOUS
- $X_n \xrightarrow{d} X \not\Rightarrow X_n - X \xrightarrow{d} 0$
  - $\hookrightarrow$ TRUE $\iff X = a = $ CONST. , WITH $p = 1$

- CONVERGENCE IN PROBABILITY :

$$X_n \xrightarrow{P} X \text{ IF } \forall \varepsilon > 0, \lim_{n\to\infty} P[|X_n - X| < \varepsilon] = 1$$

- PROPERTIES :
  - $X_n \xrightarrow{P} X \overset{\not\Leftarrow}{\Rightarrow} X_n \xrightarrow{d} X$
  - $X_n \xrightarrow{P} X \longrightarrow (X_n - X) \xrightarrow{P} 0$
- MARKOV INEQUALITY :

$$P[|X| \geq a] \leq \frac{E[|X|^k]}{a^k}, \forall a \in \mathbb{R}^+, \forall k \in \mathbb{N}^+$$

$$P[|X| < a] \geq 1 - \frac{E[|X|^k]}{a^k}$$

- COROLLARY (CHEBYCHEV'S INEQUALITY) :

$\forall$ R.V. $X \,/\, |E[X]| \cdot V[X] < \infty$:

$$P[|X - E[X]| < a] \geq 1 - \frac{V[X]}{a^2}$$

- $E[\bar{X}_n] = E[X_i] = \mu$
  - $\hookrightarrow E[S_n] = n\mu$
- $V[\bar{X}_n] = \frac{1}{n} V[X_i] = \frac{\sigma^2}{n}$
  - $\hookrightarrow V[S_n] = n\sigma^2$

- **LLN** : $(\bar{X}_n - \mu) \xrightarrow{P} 0$, $\bar{X}_n \xrightarrow{P} \mu$
- **CLT** :
  - $U_n = \dfrac{S_n - n\mu}{\sqrt{n\sigma^2}} \xrightarrow{d} Z \sim N(0,1)$, $S_n \rightsquigarrow X_1 + \dots + X_n$
  - $\bar{X}_n = \dfrac{X_1 + \dots + X_n}{n} \xrightarrow{d} N \sim \left(\mu, \dfrac{\sigma^2}{n}\right) \xrightarrow{n\to\infty} N \sim (\mu, 0)$
  - FOR $n >> 0$. $\bar{X}_n = \dfrac{S_n}{n} \sim N\left(\mu, \dfrac{\sigma^2}{n}\right) \xrightarrow{n\to\infty} N(\mu, 0)$
  - $S_n \sim N(E[S_n], V[S_n]) = N(n\mu, n\sigma^2)$

POINT ESTIMATION:

- **MME:**

eg. μ estimate for $\sim N(\ |\ )$

$$\begin{cases} \mu_1(\bar{\theta}) = M_1 = \dfrac{X_1 + \dots + X_n}{n} = E[X] \\ \mu_2(\bar{\theta}) = M_2 = \dfrac{X_1^2 + \dots + X_n^2}{n} = E[X^2] \\ \vdots \\ \mu_k(\bar{\theta}) = M_k = \dots \end{cases}$$

- $\hat{\sigma}_n = \dfrac{\sum_{i=1}^{n}(X_i - \bar{X}_n)^2}{n-1}$ CORRECTED

Asymptotic MLE: If $\hat{\theta}$ is the unbiased MLE, then as $n \to \infty$

$$\hat{\theta} \sim N\left(\theta_0, -1 / E\left(\frac{d^2}{d\theta^2} l(\theta)\right)_{\theta_0}\right)$$

Thus, C.I. of θ_0 is $\left[\hat{\theta} \pm z_{1-\alpha/2}\sqrt{(CRLB)_{\theta=\hat{\theta}}}\right]$

- **MLE:** vector of parameters

- $L(\bar{\theta} \mid \bar{x}) = f_{\bar{x},\bar{\theta}}(x_1, \dots, x_n) = \prod_{i=1}^{n} f_{\bar{\theta}}(x_i) \leadsto \hat{\theta} = f(n, x_i)$

-> compute $\dfrac{dL(\bar{\theta} \mid \bar{x})}{d\theta} = 0$, if discrete $\left(\begin{cases} f_1, & x=1 \\ \vdots \\ f_n, & x=n \end{cases}\right) \to L: f_1 \cdot \dots \cdot f_n$

- **BAYESIAN:**

- given estimators $M_1, M_2 \to V[M_1] < V[M_2] \to V[M_1]$ better

- **BIAS:** $BIAS(\hat{\theta}) = E[\hat{\theta}] - \theta$
  - IF $BIAS(\hat{\theta}) = 0 \to E[\hat{\theta}] = \theta \to \hat{\theta}$ is unbiased

- **EFFICIENCY:** IF $V[\theta_1] < V[\theta_2] \to \theta_1$ is more efficient than θ_2

- **MSE:** $MSE(\hat{\theta}) = E[(\hat{\theta} - \theta)^2] = V[\hat{\theta}] + BIAS^2(\hat{\theta})$

- **CRAMER-RAO:** $V[\hat{\theta}] \geq \left\{ n \cdot E\left[\left(\dfrac{\delta \log(f_x(b,\theta))}{\delta\theta}\right)^2\right]\right\}^{-1} =$

$\hat{\theta}$ unbiased; $\log \leadsto \ln$

$$= \left\{ -n \cdot E\left[\frac{\delta^2 \log(f_x(b,\theta))}{\delta^2 \theta^2}\right]\right\}^{-1}$$

- $\hat{\theta}_n$ CONSISTENT $\iff \hat{\theta}_n \xrightarrow{P} \theta$, $\lim_{n \to \infty} P[|\hat{\theta}_n - \theta| < \varepsilon] = 1, \forall \varepsilon > 0$
  - -> IF UNBIASED $= E[\hat{\theta}]$
  - -> CAN USE CHEBISEV INEQ.

$\hat{\theta}_n = \hat{\theta}$

- TRY USE LLN: $\bar{X}_n \longmapsto \mu = E[X]$

eg.

$$\bar{X}_n \xrightarrow{P} E[X] = f_1(\theta)$$

$$f_2(\bar{X}_n) \xrightarrow{P} f_2(E[X]) \xrightarrow[n \to \infty]{\text{CHECK IF}} \theta$$

- **INTERVAL ESTIMATION**: ($n \geq 30 \to$ LARGE, $n < 30$ SMALL)

- **I FOR μ** (σ^2 KNOWN, $n \geq 30$): / $Z = \frac{\bar{X}_n - \mu}{\sigma}$
  - LET $\alpha = 1 - p$: PROB. THAT $\mu \notin I$
  - STANDARD NORMAL
  $$\longrightarrow I(\bar{X}) = \left(\bar{X}_n - Z_{1-\frac{\alpha}{2}}\sqrt{\frac{\sigma^2}{n}},\ \bar{X}_n + Z_{1-\frac{\alpha}{2}}\sqrt{\frac{\sigma^2}{n}}\right)$$
  - UNILATERAL I:
    - $I(\bar{x}) = \left(\bar{X}_n - Z_{1-\alpha}\sqrt{\frac{\sigma^2}{n}},\ +\infty\right)$
    - $I(\bar{x}) = \left(-\infty,\ \bar{X}_n + Z_{1-\alpha}\sqrt{\frac{\sigma^2}{n}}\right)$

- FROM HIP:
  - IF σ^2 KNOWN: Z
  - IF σ^2 UNKNOWN OR $n < 30$: t_s

  - IF σ^2 UNKNOWN: $I(\bar{X})\Big|_{\sigma^2 \mapsto \hat{\sigma}^2_n}$

- $n < 30$, σ^2 UNKNOWN $\to$ **t-STUDENT**; $\rightsquigarrow \Longleftrightarrow X \sim N(\mu, \sigma^2)$
  $$T = \frac{\bar{X}_n - \mu}{\sqrt{\frac{\hat{\sigma}^2_n}{n}}} \sim t_{n-1} \quad / \ V = n-1 \text{ DOF}$$
  USABLE $\Longleftrightarrow X \sim N(\mu, \sigma^2)$
  $$\to I(\bar{x}) = \left(\bar{X}_n - t_{n-1, 1-\frac{\alpha}{2}}\sqrt{\frac{\hat{\sigma}^2_n}{n}};\ \bar{X}_n + t_{n-1,1-\frac{\alpha}{2}}\sqrt{\frac{\hat{\sigma}^2_n}{n}}\right)$$

- **I FOR σ^2**:
  - CHI-SQUARED DISTRIB.: $X = X_1^2 + \dots + X_n^2 \sim \chi^2_n$; $X_i \sim N(0,1)$
  - t-STUDENT WITH n DOF: $T_n = \frac{Z}{\sqrt{\frac{X}{n}}}$ / $Z \sim N(0,1)$, $X \sim \chi^2_n$
  - PROPERTY: MUST BE $\sim N(\cdot)$
    LET $X \sim N(\mu, \sigma^2) \to \frac{\sum_{i=1}^{n}(X_i - \bar{X}_n)^2}{\sigma^2}\left(\cdot\frac{(n-1)}{(n-1)}\right) = \frac{(n-1)}{\sigma^2}\hat{\sigma}^2_n \sim \chi^2_{n-1}$
  - $I = \left[\frac{(n-1)\hat{\sigma}^2_n}{\chi^2_{n-1,\frac{\alpha}{2}}},\ \frac{(n-1)\hat{\sigma}^2_n}{\chi^2_{n-1,1-\frac{\alpha}{2}}}\right]$ / IF UNILATERAL: $\mapsto \chi^2_{n-1,\alpha}$

- **I FOR $\mu_X - \mu_Y$** (ANY n, σ^2 KNOWN):
  - $\hat{d} = \bar{X}_{n_X} - \bar{Y}_{n_Y} \sim N\left(\mu_X - \mu_Y,\ \frac{\sigma^2_{n_X}}{n_X} + \frac{\sigma^2_{n_Y}}{n_Y}\right)$
  - $I = \left[\hat{d} - Z_{1-\frac{\alpha}{2}}\sqrt{\frac{\sigma^2_{n_X}}{n_X} + \frac{\sigma^2_{n_Y}}{n_Y}},\ \hat{d} + Z_{1-\frac{\alpha}{2}}\sqrt{\frac{\sigma^2_X}{n_X} + \frac{\sigma^2_Y}{n_Y}}\right]$

- **BAYESIAN**:
  - $1 - \alpha = P[\theta_L < \theta < \theta_U \mid X] = \int_{\theta_L}^{\theta_U} f(\theta \mid X)\, d\theta$ / $P[\theta < \theta_L \mid X] = P[\theta > \theta_U \mid X] = \frac{\alpha}{2}$

• HYPOTHESIS TESTING :

• PROB. ERROR I TYPE : $\alpha = P\left[\text{REJECT } H_0 \mid H_0 \text{ TRUE}\right]$

• PROB. ERROR II TYPE : $\beta = P\left[\text{ACCEPT } H_0 \mid H_0 \text{ FALSE}\right]$

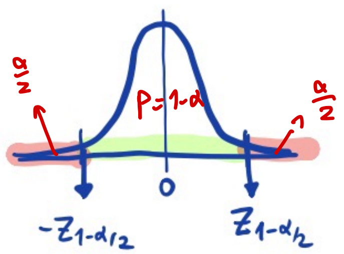
• TEST FOR μ :

• LARGE SAMPLE, σ^2 KNOWN :

USE $Z_n = \dfrac{\bar{X}_n - \mu_0}{\sqrt{\frac{\sigma^2}{n}}} \sim N(0,1)$ UNDER H_0

CHECK IF $\in$

• ACCEPTANCE REGION . $\left[-Z_{1-\frac{\alpha}{2}} , Z_{1-\frac{\alpha}{2}}\right]$



• LARGE SAMPLE, σ^2 UNKNOWN :

USE $Z_n = \dfrac{\bar{X}_n - \mu_0}{\sqrt{\frac{\hat{\sigma}_n^2}{n}}} \sim N(0,1)$ UNDER H_0

~> N.B. IF σ^2 IS ESTIMATED $(\hat{\sigma}_n)$
↳ BETTER TO USE t INSTEAD OF χ^2, BUT NOT NECESSARY

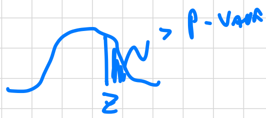
• SMALL SAMPLE, σ^2 KNOWN, $X \sim N(\mu, \sigma^2)$:

USE $Z_n = \dfrac{\bar{X}_n - \mu_0}{\sqrt{\frac{\sigma^2}{n}}} \sim N(0,1)$ UNDER H_0

• SMALL SAMPLE, σ^2 UNKNOWN, $X \sim N(\mu, \sigma^2)$:

USE $T_{n-1} = \dfrac{\bar{X}_n - \mu_0}{\sqrt{\frac{\hat{\sigma}_n^2}{n}}} \sim t_{n-1}$ UNDER H_0

• ACCEPTANCE REGION . $\left[-t_{n-1, 1-\frac{\alpha}{2}} ; t_{n-1, 1-\frac{\alpha}{2}}\right]$



• p - VALUE :

AREA ON RIGHT OF $Z = \dfrac{\bar{X}_n - \mu_0}{\sqrt{\frac{\sigma^2}{n}}}$

(CONSIDER $X \sim N(\ )$)

↳ p-VALUE < $\frac{\alpha}{2}$ (OR α) -> H_0 REJECTED
p-VALUE > $\frac{\alpha}{2}$ (OR α) -> H_0 ACCEPTED

• TEST FOR σ^2 $(X \sim N(\mu, \sigma^2))$:

• PROPERTY : $Q_{n-1} = \dfrac{(n-1)\hat{\sigma}_n^2}{\sigma_0^2} \sim \chi^2_{n-1}$

• ACCEPTANCE REGION = $\left[\chi^2_{n-1, \frac{\alpha}{2}} ; \chi^2_{n-1, 1-\frac{\alpha}{2}}\right]$

-> IF $n_x, n_y < 30$, σ^2 KNOWN -> Z

UNILATERAL : $t_{n-1, 1-\alpha}$

↑ AND σ^2 UNKNOWN

• TEST FOR $\mu_X - \mu_Y$ $(n_x, n_y \geq 30)$:

• $\bar{X}_{n_1} - \bar{Y}_{n_2} \sim N\left(\overbrace{\mu_X - \mu_Y}^{d_0}, \frac{\sigma_X^2}{n_1} + \frac{\sigma_Y^2}{n_2}\right)$

-> COMPUTE $D_{n_1, n_2} = \dfrac{(\bar{X}_{n_1} - \bar{Y}_{n_2}) - d_0}{\sqrt{\frac{\sigma_X^2}{n} + \frac{\sigma_Y^2}{n}}} \sim N(0,1)$

↳ / H_0 TRUE

• ACCEPTANCE REGION = $\left[-Z_{1-\frac{\alpha}{2}} ; Z_{1-\frac{\alpha}{2}}\right]$

• TEST FOR $\mu_X - \mu_Y$ $(n_x, n_y < 30)$:

• ESTIMATOR OF σ^2 : $\hat{\sigma}_m^2 = \dfrac{(n_x - 1)\hat{\sigma}_X^2 + (n_y - 1)\hat{\sigma}_Y^2}{n_x + n_y - 2}$

• PROPERTY : $T = \dfrac{(\bar{X}_{n_x} - \bar{Y}_{n_y}) - d_0}{\sqrt{\frac{(n_x + n_y)}{n_x n_y} \cdot \hat{\sigma}_m^2}} \sim t_{n_x + n_y - 2}$

↳ / UNDER H_0 TRUE

• ACCEPTANCE REGION . $\left[-t_{n_x + n_y - 2, 1-\frac{\alpha}{2}} ; t_{n_x + n_y - 2, 1-\frac{\alpha}{2}}\right]$

• LIKELIHOOD RATIO TEST:

GIVEN $H_0: \hat{\theta} \in \Theta_0$, $H_1: \hat{\theta} \in \bar{\Theta}_0$ $\rightarrow \lambda(\bar{x}) = \dfrac{\sup_{\Theta_0} L(\theta|\bar{x})}{\sup_{\bar{\Theta}_0} L(\theta|\bar{x})}$

MAX LIKELIHOOD ESTIMATOR

• $\alpha = P[\underbrace{\lambda(\bar{x}) < \lambda^*}_{\text{REJECTION REGION}} \mid H_0 \text{ IS TRUE}] \rightsquigarrow \lambda^* = \dots$

↳ FIND $\bar{x} \lessgtr \dots$

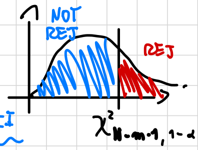
$\rightarrow H_0$ REJECTED IF $\lambda(\bar{x}) < \lambda^*$

• BAYESIAN TESTS:

H_0 REJECT $\iff \dfrac{P(\hat{\theta} \in \Theta_0 \mid \bar{x})}{P(\hat{\theta} \in \Theta_1 \mid \bar{x})} < c$ / IF $c = 1$ ↳ MAP TEST

• GOODNESS OF FIT: ⇝ TEST FOR A DISTRIBUTION

• $H_0: X \sim f = f_0$, $H_1: X \sim f \neq f_0$



• χ^2 GOF:

• THEOREM, $W = \sum_{i=1}^{N} \dfrac{(f_i - mp_i)^2}{mp_i} \sim \chi^2_{N-m-1,\,\underset{\sim}{1-\alpha}}$

f_i: # OUTCOME I_i OCCURRED

IF $mp_i \geq 5, \forall i$

/ $p_i = P[X \in I_i \mid f_0$ TRUE DENSITY$]$

CI

$\rightarrow$ IF $W < \chi^2_{m-1,\,1-\alpha}$ $\rightarrow H_0$ NOT REJECTED

• SAME CAN BE APPLIED IF f_i FROM A SAMPLE $X = [X_1, \dots, X_n]$

$\rightarrow$ CHECK $W < \chi^2_{N-m-1,\,1-\alpha}$ / N: N° INTERVALS

m ⇝ # PARAM. ESTIMATED THROUGH THE SAMPLE

NO EXAM

• KOLMOGOROV-SMIRNOV GOF TEST: ⇝ (SMALL/LARGE SAMPLES, f_X ABS. CONTINUOUS)

• IT COMPARES $\hat{F}(x)$ AND $F_0(x)$

$\hat{F}$ ⇝ EMPIRICAL, F_0 ⇝ THEORETICAL

• MEASURE: $D = \sup_{x \in \mathbb{R}} |\hat{F}(x) - F_0(x)| = D = \dots$ (TABLES)

• THEOREM: $\forall d > 0 \rightarrow P(D \geq d)$ c/ F_0, c.) n (IF $X \sim F_0$)

• $P[D \geq d^*_{1-\alpha}] = \alpha$, REJ REGION: $\bar{C} = (d^*_{1-\alpha}, 1]$

• TEST FOR INDEPENDENCE:

TEST H_0: X, Y INDEPENDENT

• $W = \sum_{i=1}^{N_X} \sum_{j=1}^{N_Y} \dfrac{(f_{ij} - n\hat{p}_i\hat{q}_j)^2}{n\hat{p}_i\,\hat{q}_j} \sim \chi^2_{(N_X-1)(N_Y-1),\,1-\alpha}$

IF $n\hat{p}_i\hat{q}_j \geq 5, \forall i,j$

• CHECK H_0 AS IN χ^2 GOF

/ $\hat{p}_i$ = OBSERVED $P[X \in I^X_i]$, $\hat{q}_j$ ANALOG...

- CONFIDENCE INTERVAL FOR $Exp(p)$:

LET $X_i \sim Exp(\lambda)$ / $Y = \sum X_i \sim \Gamma(n, \lambda)$

- PROPERTY: $2\lambda Y \sim \chi^2_{2n}$

$$\rightarrow 1-\alpha = P\left[\chi^2_{2n,\frac{\alpha}{2}} \leq 2\lambda Y \leq \chi^2_{2n,1-\frac{\alpha}{2}}\right]$$

$$\rightarrow \underset{\text{FOR } \lambda}{I} = \left[\frac{\chi^2_{2n,\frac{\alpha}{2}}}{2 \cdot \sum X_i} ; \frac{\chi^2_{2n,1-\frac{\alpha}{2}}}{2\sum X_i}\right]$$

- CONFIDENCE INTERVAL FOR $Bernoulli(p)$:

LET $X_i \sim Bernoulli(p)$ / $Y = \sum X_i \sim Bin(n, p)$

$$\rightarrow 1-\alpha = P\left[-Z_{1-\alpha/2} \leq \frac{Y - np}{\sqrt{np(1-p)}} \leq Z_{1-\frac{\alpha}{2}}\right]$$

/ $\hat{p} = E[X_i] = \frac{\bar{X}_n}{n}$

$$\rightarrow I = \left[\hat{p} - Z_{1-\frac{\alpha}{2}} \cdot \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} ; \hat{p} + Z_{1-\frac{\alpha}{2}} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\right]$$

- DELTA METHOD:

USED TO ESTIMATE A f OF A PARAMETER / $Y = \sum X_i$

- LET $g: \mathbb{R} \mapsto \mathbb{R}$ / CONTINOUS $g'(\theta)$, $g'(\theta) \neq 0$

$$\rightarrow \sqrt{n}\left(g(Y) - g(\overset{\mu}{\overset{\shortparallel}{\theta}})\right) \xmapsto{d} N\left(0, \sigma^2 \left[g'(\theta)\right]^2\right)$$

$$\xrightarrow{CLT} g(\bar{X}_n) \xmapsto{d} N\left(g(\mu), \frac{\sigma^2 [g'(\theta)]^2}{n}\right)$$

$$\rightarrow 1-\alpha = P\left[-Z_{1-\frac{\alpha}{2}} \leq \frac{g(\bar{X}_n) - g(\mu)}{\sqrt{\frac{\sigma^2[g'(\theta)]^2}{n}}} \leq Z_{1-\frac{\alpha}{2}}\right]$$