CONFIDENCE INTERVALS

- RELIES A LOT ON QUANTILI
- STD. SAMPLE MEAN: $Z = \frac{\bar{X} - \mu}{\sigma / \sqrt{N}}$
- AL 95% $P\left(-1.96 < \frac{\bar{X} - \mu}{\sigma / \sqrt{N}} < 1.96\right) = .95$
- HIGHER CONFIDENCE $\longrightarrow$ WIDER INTERV
- C.I. $\left(\bar{X} - 1.96 \frac{\sigma}{\sqrt{N}}, \bar{X} + 1.96 \frac{\sigma}{\sqrt{N}}\right)$ PROBABILITY THAT μ IS WITHIN INTERVAL IS 95%
- $\bar{X} \pm Z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{N}}$ **GENERAL FORM**
- **SAMPLE SIZE:** $N = \left(2 Z_{\alpha/2} \cdot \frac{\sigma}{W}\right)^2$
- **CAN BOOTSTRAP**
- **CONFIDENCE LEVEL:** $\gamma = 2\,\mathrm{ERF}\left(z = \frac{\sqrt{N}}{\sigma}\delta\right)$, δ SCARTO
- **VARIANCE UNKNOWN**, SE $N > 40$ MAKES SENSE TO $\bar{X} \pm Z_{\alpha/2} \frac{S}{\sqrt{N}}$. S IS SAMPLE VARIANCE

- **FOR A GENERAL θ**

$\hat{\theta} \pm (Z \text{ CRIT. VALUE})(\text{EST. STD. ERROR OF } \hat{\theta})$

$\hat{\theta} \pm Z_{\alpha/2} \cdot \hat{\sigma}_\theta$

- **POPULATION PROPORTION**
- $P = \hat{P} \pm Z_{\alpha/2} \frac{\sqrt{\hat{P}(1-\hat{P})/N + Z_{\alpha/2}^2/4N^2}}{1 + Z_{\alpha/2}^2/N}$ BINOMIAL ASSUMPTIONS
- **IF $N \gg$** $\hat{P} \pm Z_{\alpha/2} \sqrt{\hat{P}\hat{Q}/N}$

- **CONFIDENCE BOUNDS (ONE SIDED)**

REPLACE $Z_{\alpha/2}$ WITH Z_α AND TAKE IT ONE SIDED

$\mu < \bar{X} + Z_\alpha \frac{\sigma}{\sqrt{N}}$; $\mu > \bar{X} - Z_\alpha \frac{\sigma}{\sqrt{N}}$

- **DRAWING FROM NORMAL DISTRIBUTIONS**, μ, σ UNKNOWN

SE $N \gg \longrightarrow$ OK

SE $N \ll \rightarrow$ $T = \frac{\bar{X} - \mu}{S / \sqrt{N}}$, μ NOTO, T IS t-DISTRIBUTION A $N-1$ DOF. T CRIT VALUE IS $t_{z,\alpha}$

C.I. $\bar{X} \pm t_{\alpha/2, N-1} \cdot \frac{S}{\sqrt{N}}$

PREDICTION INTERVAL FOR PREDICTION $\bar{X} \pm t_{\alpha/2, N-1} \cdot S\sqrt{1 + \frac{1}{N}}$ | $t = \frac{\bar{X} - X_{N+1}}{S\sqrt{1 + 1/N}}$

TODO TOLERANCE INTERVALS

- **CI FOR VARIANCE, STDDEV**

$\frac{\Sigma(X_i - \bar{X})}{\sigma^2}$ IS χ^2 A $N-1$ DOF. NOT SYMMETRIC!!! $\rightarrow$ $\frac{(N-1)S^2}{\chi^2_{\alpha/2; N-1}} < \sigma^2 < \frac{(N-1)S^2}{\chi^2_{1-\alpha, N-1}}$; PER σ PRENDO SQRT

- **LARGE SAMPLES** $\frac{\hat{S}_N^2}{1 + \sqrt{\frac{2}{N}} Z_\alpha} \leq \sigma^2 \leq \frac{S_N^2}{1 - \sqrt{\frac{2}{N}} Z_\alpha}$

GENERIC CASE, LARGE SAMPLES

θ, $\hat{\theta}$ = MLE, $f(x,\theta)$, $I(\theta)$

$\theta = \hat{\theta} \pm z_{\alpha/2} \cdot \frac{\hat{\theta}'}{\sqrt{n I(\theta)}}$;

Point Estimation

$\hat{\theta}$ P.E. di qualcosa θ, $\hat{\theta} = \theta + \text{error}$, $E[(\hat{\theta}-\theta)^2] = \text{VAR}(\theta) + [E(\theta) - \theta]^2$ **VAR + BIAS!**

- $\hat{\theta}$ consistente se bias $\to 0 \mapsto N \to \infty$
- $\hat{\theta}$ unbiased if $E(\hat{\theta}) = \theta \;\forall \theta$, biased < unbiased < MVUE
- prob of $\hat{\theta}$ centered on θ; in Gauss $\bar{\mu}$ is MVUE
- **Efficiency:** $e(\hat{\theta}) = \frac{I^{-1}(\theta)}{\text{VAR}(\hat{\theta})}$

Sample variance:

$$\hat{\sigma}^2 = \frac{\sum (x_i - \bar{x})^2}{N-1}$$

, unbiased for σ^2; se divido per N ho bias

STD error: $\sqrt{\text{VAR}(\hat{\theta})}$ **Sample std. error:** $\sqrt{\frac{\text{VAR}(\hat{\theta})}{s}}$

Bootstrapping: simulo e calcolo $\hat{\theta}_N \to$ sample std. dev. of runs

Method of Moments

Determino parametri distribuzione.
- uguaglio momenti sample a momenti popolazione

Sample moments: $\frac{1}{N}\sum_1^N x_i^K$
- risolvo x parametri

Esempio (Gamma)

$$\begin{cases} E(x) = \alpha\beta, \quad E(x^2) = \frac{\beta^2\Gamma(\alpha+2)}{\Gamma(\alpha)} = \beta^2(\alpha+1)\alpha \\ \bar{x} = \alpha\beta \quad \frac{1}{N}\sum x^2 = \alpha(\alpha+1)\beta^2 \\ \hat{\alpha} = \frac{\bar{x}^2}{(1/N)\sum x^2 - \bar{x}^2} \quad \hat{\beta} = \frac{(1/N)\sum x^2 - \bar{x}^2}{\bar{x}} \end{cases}$$

Maximum Likelihood Estimation

Requires independence!

Likelihood fcn: joint PMF: $f_{x_1} \cdot f_{x_2} \cdots f_{x_N}$

Massimizzo $\log[L]$ x' prodotti $\to$ somme

- $$\frac{d}{d(\theta_1 \ldots \theta_N)} \ln\left[f(x_1 \ldots x_n; \theta_1 \ldots \theta_N)\right] = 0$$
- most probable Bayesian estimator
- se $\theta_N > 1$, derivate parziali e risolvi sistema
- for large N, MLE's are MVUE

- **Invariance principle:** MLE of $f(\theta_1 \ldots \theta_N) \to f(\hat{\theta}_1 \ldots \hat{\theta}_N)$
- **Asimptotic normality:** $N \to \infty \mid \text{MLE} \to N \sim (\theta, I^{-1})$, $I^{-1} =$ Fisher info matrix^{-1} $\to$ Condizioni: $\begin{cases} \text{log-likelyhood 2-diff.} \\ I \neq 0, \text{ continua} \\ \text{MLE consistente} \end{cases}$
- **Efficiency:** $N \to \infty$, Cramer-Rao bound: no other consistent estimator has lower asymptotic MSE
- **CR bound:**
  - limit on variance of unbiased estimator: $\text{VAR}(\hat{\theta}) \geq I^{-1}(\theta)$, bound is I^{-1}
  - multivariate $$\text{cov}(\mathbf{T}(x)) \geq \frac{\partial \psi}{\partial \boldsymbol{\theta}} [I^{-1}(\boldsymbol{\theta})] \left(\frac{\partial \boldsymbol{\psi}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}\right)^T$$, $\frac{\partial \boldsymbol{\psi}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}$ is **Jacobian**, $\frac{\partial \psi_i(\theta)}{\partial \theta_j}$
  - multivariate + unbiased $\text{cov}(\mathbf{T}(x)) \geq I^{-1}(\boldsymbol{\theta})$
  - es. Gaussiana

Fisher Information

- measures info of X (RV) carries about θ (par).
- variance of score
- $I_n(\theta) = n \cdot I(\theta)$

in n samples ho n times $I(\theta)$ of single sample

- 1st derivative of log-likelihood

1st moment

- **Score:** $\frac{\partial}{\partial \theta} \log\left[L(x_1 \dots x_n, \theta_1 \dots \theta_n) \mid \theta\right]$, $E[\text{score}] = 0$
- **F.I.:** $I(\theta) = E\left[\left(\frac{\partial}{\partial \theta} \log\left[L(x_1 \dots x_n, \theta_1 \dots \theta_n)\right]\right)^2\right] = \int \left(\frac{\partial}{\partial \theta} \log\left[L(x,\theta)\right]\right)^2 f(x,\theta)\, dx$

(2 moment of score)

- **F.I.:** if $L(x,\theta)$ is 2-diff: $I(\theta) = -E\left[\frac{\partial^2}{\partial \theta^2} \log\left[L(x,\theta)\right] \Big| \theta\right]$ under regularity con.
- **Multivariate: F.I. matrix**
- $\theta = [\theta_1 \dots \theta_n]^T$, $N \times N$ $\qquad T(x) = [T_1(x) \dots T_n(x)]$ estimator vector, $E[T(x)] = \psi(\theta)$
- $I_{m,k} = -E\left[\frac{\partial^2}{\partial \theta_m \partial \theta_k} \cdot \log L(x,\theta)\right] =$ expect. of Hessian of relative entropy (curvature)

↳ or K.L. divergence

- $I = \mathrm{cov}\left[\frac{\partial \log L(\theta)}{\partial \theta}\right]$ under regularity conditions

Statistica: v.a. T funzione del campione

Monte Carlo Estimation

Because it may be hard to integrate, change variables, etc...

Approx $E(x)$ of any function of RV

$$E[f(x)] = \int f(x) p(x)\, dx \approx \frac{1}{S} \sum^{S} f(x_s)$$

- change var by applying f to drawn samples

θ ψ θ $\qquad$ θ Ψ θ $\qquad$ θ ψ θ $\qquad$ θ $\qquad$ ψ θ $\qquad$ T

θ $\quad$ I $\quad$ T