ENGG 2430 / ESTR 2004: Probability and Statistics

Spring 2019

12. Classical statistics

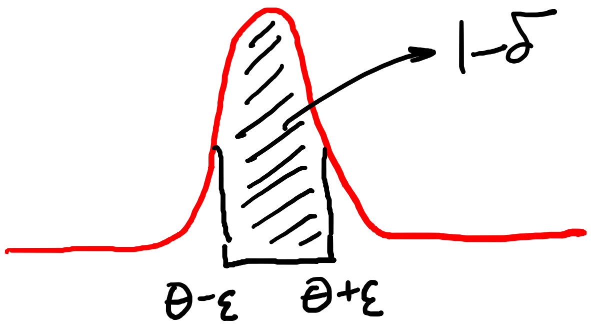
Andrej Bogdanov

Estimators

$\boldsymbol{X} = (X_1, \ldots, X_n)$ independent samples

Unbiased: $\mathbf{E}[\hat{\Theta}_n] = \theta$

Consistent: $\hat{\Theta}_n$ converges to θ in probability



$\forall \varepsilon, \delta: \lim_{n \to \infty} P(|\hat{\Theta}_n - \theta| > \varepsilon) < \delta$

Estimating the mean

$\boldsymbol{X} = (X_1, \ldots, X_n)$ **independent samples of** X

$$\hat{M} = (X_1 + \ldots + X_n) \;/\; n$$

Unbiased? $E[M] = E\left[\frac{X_1 + \ldots + X_n}{n}\right] = \frac{1}{n}\left(E[X_1] + \ldots + E[X_n]\right) = E[X]$

YES.

Consistent? YES, BY WEAK LAW OF LARGE NUMBERS

Maximum likelihood

Bayesian MAP estimate:

maximize $f_{\Theta \mid \boldsymbol{X}}(\theta \mid \boldsymbol{x}) = f_{\boldsymbol{X} \mid \Theta}(\boldsymbol{x} \mid \theta)\, f_{\Theta}(\theta)$

Classical ML (maximum likelihood) estimate:

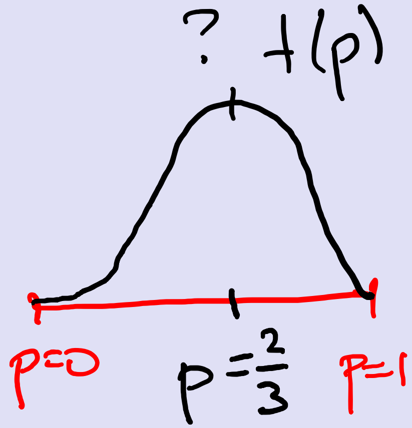
maximize $f_{\boldsymbol{X} \mid \Theta}(\boldsymbol{x} \mid \theta)$

Coin flip sequence HHT. What is ML bias estimate?

PARAMETER: p

$$P(HHT \mid p) = p^2(1-p)$$

WHICH p MAXIMIZES $f(p) = p^2(1-p)$?



$$\frac{df(p)}{dp} = 2p(1-p) - p^2$$

$$\frac{df(p)}{dp} = 0 \longrightarrow 2p(1-p) = p^2$$

$$2(1-p) = p$$

$$p = \frac{2}{3}$$

ML = MAP WITH UNIFORM PRIORS

Maximum likelihood for $\mathrm{Bernoulli}(p)$

k heads, $n-k$ tails. ML bias estimate?

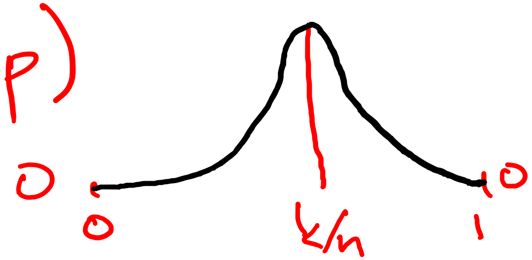
$$P(\underbrace{k \text{ HEADS, } n-k \text{ TAILS}}_{\text{IN A SPECIFIC ORDER}} \mid p) = p^k (1-p)^{n-k}$$

FIND p THAT MAXIMIZES $f(p) = p^k (1-p)^{n-k}$

$$\frac{d}{dp} f(p) = k p^{k-1} (1-p)^{n-k} - (n-k) p^k (1-p)^{n-k-1}$$

$$= p^{k-1} (1-p)^{n-k-1} \left(k(1-p) - (n-k)p \right)$$

$= 0$ WHEN $p=0$ OR $p=1$ OR $p = \frac{k}{n}$



ML ESTIMATE IS $\hat{p} = \frac{k}{n}$.

UNBIASED, CONSISTENT

Within the first 3 seconds, raindrops arrive at times 1.2, 1.9, and 2.5. What is the estimated rate?

RATE = λ

MODEL: Poisson(λ) DROP IN EACH sec (TIME UNIT = 1 sec)

$$P(X_1=0, X_2=2, X_3=1 \mid \lambda) = P(X_1=0\mid\lambda)\,P(X_2=2\mid\lambda)\,P(X_3=1\mid\lambda)$$

$$= e^{-\lambda}\cdot\frac{\lambda^0}{0!}\cdot e^{-\lambda}\cdot\frac{\lambda^2}{2!}\cdot e^{-\lambda}\cdot\frac{\lambda^1}{1!}$$

$$f(\lambda) \qquad = \frac{e^{-3\lambda}\cdot\lambda^3}{2}$$

$$\frac{df(\lambda)}{d\lambda} = \frac{-3e^{-3\lambda}\lambda^3 + 3e^{-3\lambda}\cdot\lambda^2}{2} = \frac{3}{2}e^{-3\lambda}\lambda^2(1-\lambda)$$

ZERO AT $\lambda=0$, $\lambda=1$

$$\boxed{\hat{\Lambda} = 1}$$

Within the first 3 seconds, raindrops arrive at times 1.2, 1.9, and 2.5. What is the estimated rate?

UNIT OF TIME = 3 sec

DROPS IN FIRST 3 sec IS Poisson(λ) R.V. X

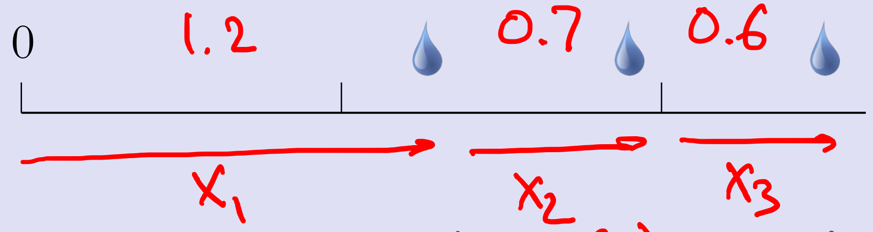
$$P(X=3 \mid \lambda) = e^{-\lambda} \cdot \frac{\lambda^3}{3!}$$

$$\frac{df(\lambda)}{d\lambda} = \frac{1}{3!}\left(-\lambda e^{-\lambda}\lambda^3 + 3e^{-\lambda}\lambda^2\right)$$

ZERO WHEN $\lambda=0$ OR $\lambda = 3$

$$\boxed{\hat{\lambda} = 3}$$

The first 3 raindrops arrive at 1.2, 1.9, and 2.5 sec. What is the estimated rate?



X_1, X_2, X_3 IND Exponential(λ) R.V.S

$$f_{X_1X_2X_3}(1.2, 0.7, 0.6 \mid \lambda) = \lambda e^{-1.2\lambda} \cdot \lambda e^{-0.7\lambda} \cdot \lambda e^{-0.6\lambda}$$

$$= \lambda^3 e^{-2.5\lambda}$$

$$\frac{df(\lambda)}{d\lambda} = 3\lambda^2 e^{-2.5\lambda} - 2.5\lambda^3 e^{-2.5\lambda}$$

$$= e^{-2.5\lambda}\lambda^2 \underline{(3 - 2.5\lambda)}.$$

$= 0$ WHEN $\lambda = \frac{3}{2.5}$

$$\boxed{\hat{\lambda} = 3/2.5}$$

Maximum likelihood for $\mathrm{Exponential}(\lambda)$

OBSERVE FIRST n SAMPLES AT TIMES $T_1, \ldots, T_n$

ML ESTIMATE FOR λ IS $\boxed{\hat{\lambda} = \frac{n}{T_n}}$

NOT UNBIASED!

$$E\left[\frac{n}{T_n}\right] = E\left[\frac{n}{X_1 + \ldots + X_n}\right]$$

$$\lambda = \frac{n}{E[X_1 + \ldots + X_n]}$$

NOT EQUAL

CONSISTENT: EVENTUALLY T_n CLOSE TO $\frac{1}{\lambda} \cdot n$

A $\mathrm{Normal}(\mu, \sigma)$ **RV takes values** $2.9, 3.3$**. What is the ML estimate for** μ**?**

$$f(2.9, 3.3 \mid \mu) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(2.9-\mu)^2/2\sigma^2} \cdot \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(3.3-\mu)^2/2\sigma^2}$$

MAXIMIZING f IS SAME AS MAXIMIZING $\ln f$.

$$\ln f = -\frac{(2.9-\mu)^2}{2\sigma^2} - \frac{(3.3-\mu)^2}{2\sigma^2} + C$$

$$\frac{\partial(\ln f)(\mu)}{\partial \mu} = \frac{1}{\sigma^2}\left((2.9-\mu) + (3.3-\mu)\right)$$

ZERO WHEN $\mu = \frac{2.9+3.3}{2}$

$$\boxed{\hat{\mu} = \frac{2.9+3.3}{2} = 3.1}$$

A $\mathrm{Normal}(\mu, \sigma)$ **RV takes values** $2.9, 3.3$**. What is the ML estimate for** $v = \sigma^2$**?**

$$f(2.9, 3.3 \mid v) = \frac{1}{\sqrt{2\pi v}} e^{-(2.9-\mu)^2/2v} \cdot \frac{1}{\sqrt{2\pi v}} e^{-(3.3-\mu)^2/2v}$$

$$\ln f = \ln\left(\frac{1}{v}\right) - \frac{(2.9-\mu)^2 + (3.3-\mu)^2}{2v} + C$$

$$\frac{\partial(\ln f)(v)}{\partial v} = -\frac{1}{v} + \frac{(2.9-\mu)^2 + (3.3-\mu)^2}{2v^2}$$

ZERO WHEN $\hat{v} = \dfrac{(2.9-\mu)^2 + (3.3-\mu)^2}{2}$

$$\left.\begin{array}{l} \dfrac{\partial(\ln f)(\hat{\mu}, \hat{v})}{\partial \mu} = 0 \\[2ex] \dfrac{\partial(\ln f)(\hat{\mu}, \hat{v})}{\partial v} = 0 \end{array}\right\} \text{ AT } \begin{array}{l} \hat{\mu} = \dfrac{2.9 + 3.3}{2} = 3.1 \\[2ex] \hat{v} = \dfrac{(2.9-\hat{\mu})^2 + (3.3-\hat{\mu})^2}{2} = 0.04 \end{array}$$

Maximum likelihood for $\mathrm{Normal}(\mu, \sigma)$

$(X_1, \ldots, X_n)$ **independent** $\mathrm{Normal}(\mu, \sigma)$

Joint ML estimate $(\hat{M}, \hat{V})$ **of** $(\mu, v = \sigma^2)$**:**

$$\hat{M} = \frac{X_1 + \ldots + X_n}{n}$$

UNBIASED

$$\hat{V} = \frac{(X_1 - \hat{M})^2 + \ldots + (X_n - \hat{M})^2}{n}$$

?

$$\mathbf{E}[\hat{V}] = E\left[\frac{(X_1-\hat{M})^2+\dots+(X_n-\hat{M})^2}{n}\right]$$

$$= E\left[\frac{X_1^2+\dots+X_n^2}{n}\right] - E\left[\left(\frac{X_1+\dots+X_n}{n}\right)^2\right]$$

$$= \frac{\sum E[X_i^2]}{n} - \frac{\sum E[X_i^2] + \sum_{i\neq j} E[X_iX_j]}{n^2}$$

$$= \left(1-\frac{1}{n}\right)\frac{\sum E[X_i^2]}{n} - \frac{\sum_{i\neq j} E[X_i]E[X_j]}{n^2}$$

$$= \left(1-\frac{1}{n}\right)E[X^2] - \frac{n(n-1)\,E[X]^2}{n^2}$$

$$= \frac{n-1}{n}\left(E[X^2]-E[X]^2\right)$$

$$= \frac{n-1}{n}\cdot v$$

NOT UNBIASED!

$(X_1, \ldots, X_n)$ **independent** Normal(μ, σ)

$$\hat{M} = \frac{X_1 + \ldots + X_n}{n}$$

UNBIASED, CONSISTENT ESTIMATOR OF μ

$$\frac{n}{n-1}\hat{V} = \frac{(X_1 - \hat{M})^2 + \ldots + (X_n - \hat{M})^2}{n-1}$$

UNBIASED, CONSISTENT ESTIMATOR OF $v = \sigma^2$

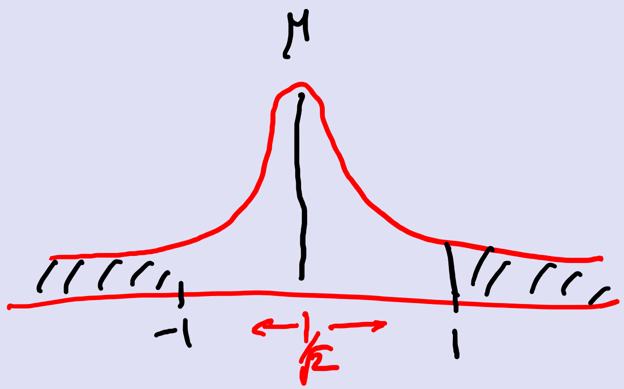
A $\mathrm{Normal}(\mu, 1)$ **RV takes values** X_1, X_2**. You estimate the mean by** $\hat{M} = (X_1 + X_2)/2$**. What is the probability that** $|\hat{M} - \mu| > 1$**?**

X_1, X_2 : Normal$(\mu, 1)$

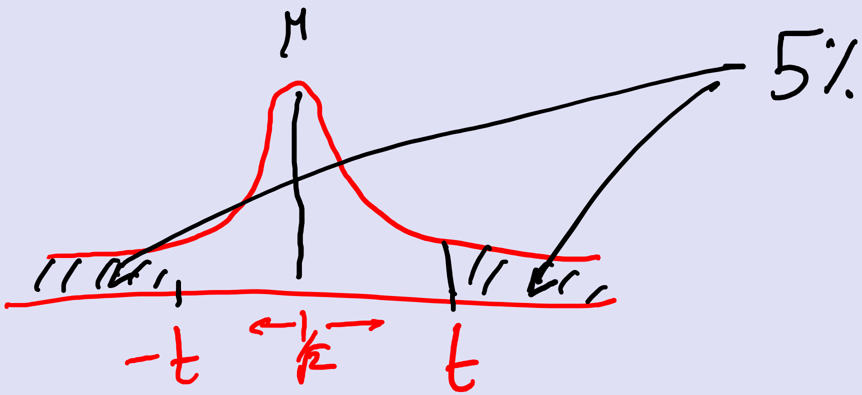
$$\mathrm{Var}\left[\frac{X_1 + X_2}{2}\right] = \frac{1}{4}\left(\mathrm{Var}[X_1] + \mathrm{Var}[X_2]\right) = \frac{1}{2}$$

$$\hat{M} = \frac{X_1 + X_2}{2} \quad : \quad \text{Normal}(\mu, 1/\sqrt{2})$$

$$P(|\hat{M} - \mu| > 1) = 2\,P(\text{Normal}(0,1) > \sqrt{2})$$
$$\approx 2 \cdot 0.079$$
$$= 15.8\%$$



For which value of t can we guarantee $|\hat{M} - \mu| \leq t$ with 95% probability?



$P(\text{Normal}(0,1) \leq -n) = 0.025$

FOR $n \approx 1.960$

$$P(|\hat{M} - \mu| \leq -t) = P(\text{Normal}(0,1) \leq -\sqrt{2}\,t) = 0.025$$

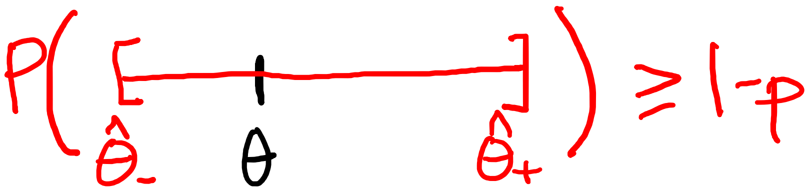
WHEN $t \approx \frac{1.960}{\sqrt{2}} \approx 1.386$

$\left(\frac{X_1 + X_2}{2} - 1.386,\ \frac{X_1 + X_2}{2} + 1.386\right)$ IS A $\boxed{\text{95\%-CONFIDENCE INTERVAL}}$

Confidence intervals

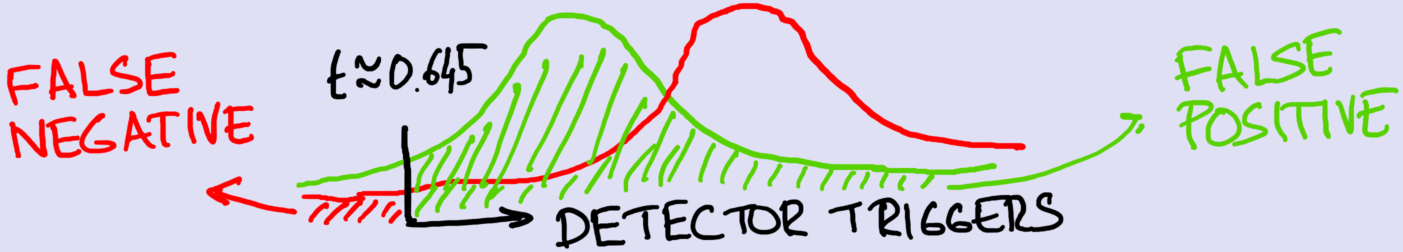
A p-**confidence interval** is a pair $\hat{\Theta}_-$, $\hat{\Theta}_+$ so that

$$\mathbf{P}(\theta \text{ is between } \hat{\Theta}_- \text{ and } \hat{\Theta}_+) \geq p$$



An car-jack detector outputs $\mathrm{Normal}(0, 1)$ if there is no intruder and $\mathrm{Normal}(1, 1)$ if there is.

You want to catch 95% of intrusions. What is the probability of a false positive?



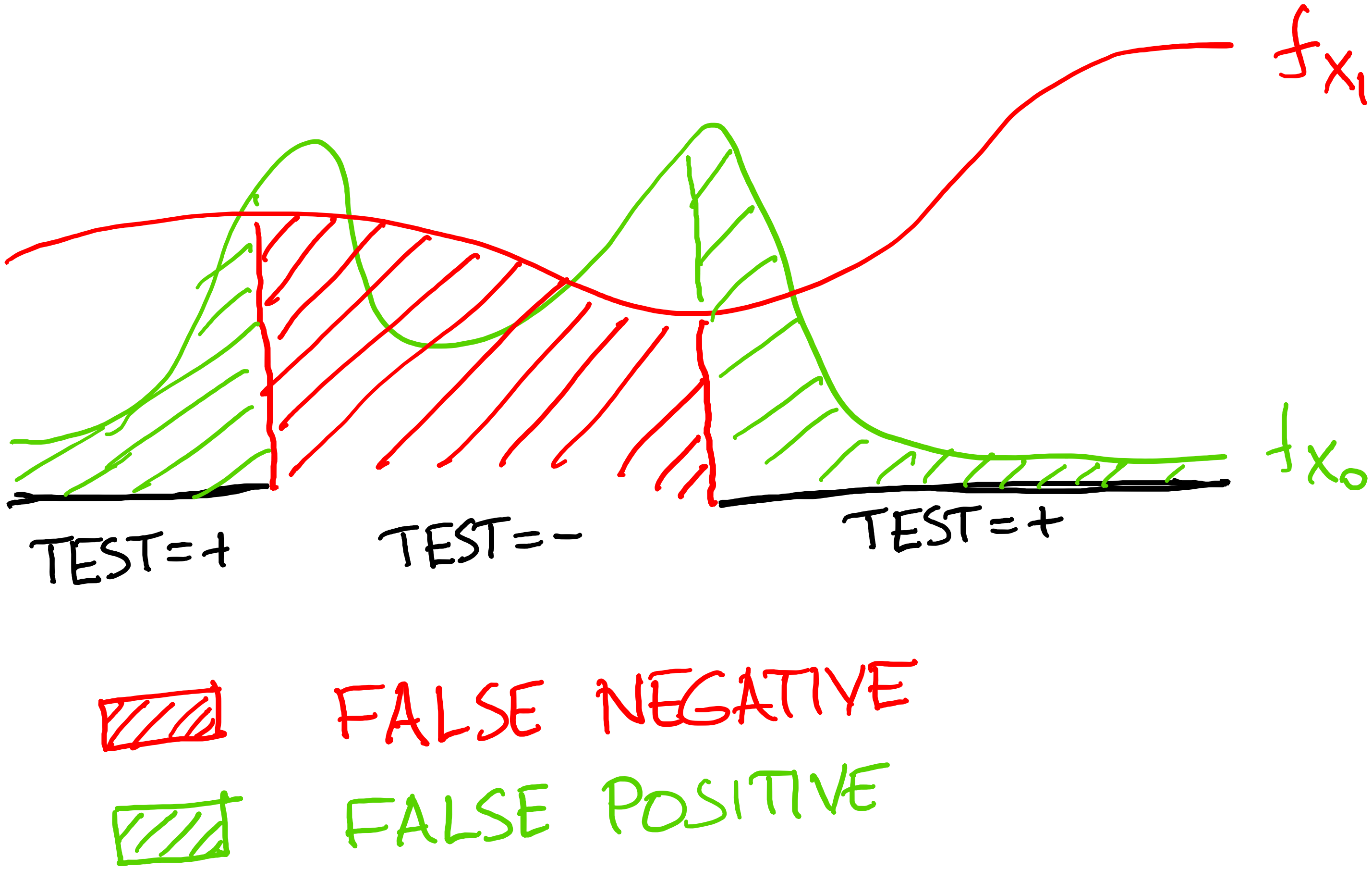
INTRUDER: $X_1 = \mathrm{Normal}(1,1)$

FALSE NEGATIVE: $P(X_1 < t) = 5\% \longrightarrow t = -0.645$

LEGIT: $X_0 = \mathrm{Normal}(0,1)$

FALSE POSITIVE: $P(X_0 > t) \approx 73\%$

Hypothesis testing



Neyman-Pearson Lemma

Among all X_1/X_0 tests with given false negative probability, the false positive is minimized by the one that picks samples with largest likelihood ratio

$$\frac{f_{X_1}(x)}{f_{X_0}(x)}$$

Rain usually falls at 1 drop/sec. You want to test today's rate is 5/sec based on first drop. How to set up test with 5% false negative?

$f_1(x) = 5e^{-5x}$ Exponential(5)

$f_0(x) = e^{-x}$ Exponential(1)

$\left.\right\} \dfrac{f_1(x)}{f_0(x)} = 5e^{-4x}$ DECREASES IN x

NEYMAN-PEARSON: TEST $= +$ FOR $x \in [0, t]$

WHERE $P(\text{Exponential}(5) > t) = 5\%$ (FALSE NEG.)

$e^{-5t} = 0.05 \longrightarrow t \approx 0.60$

$P(\text{Exponential}(1) < t) \approx 1 - e^{-0.60} \approx 45\%$ (FALSE POS.)