A short course on

Probabilities and Random Variables

This video:

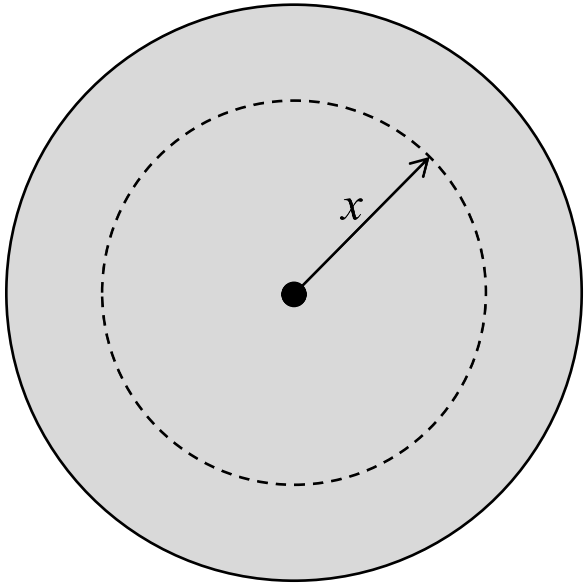
Classical & Bayesian Statistical Inference

Terje's Toolbox is freely available at terje.civil.ubc.ca
It is created and maintained by Professor Terje Haukaas, Ph.D., P.Eng.,
Department of Civil Engineering, The University of British Columbia (UBC), Vancouver, Canada

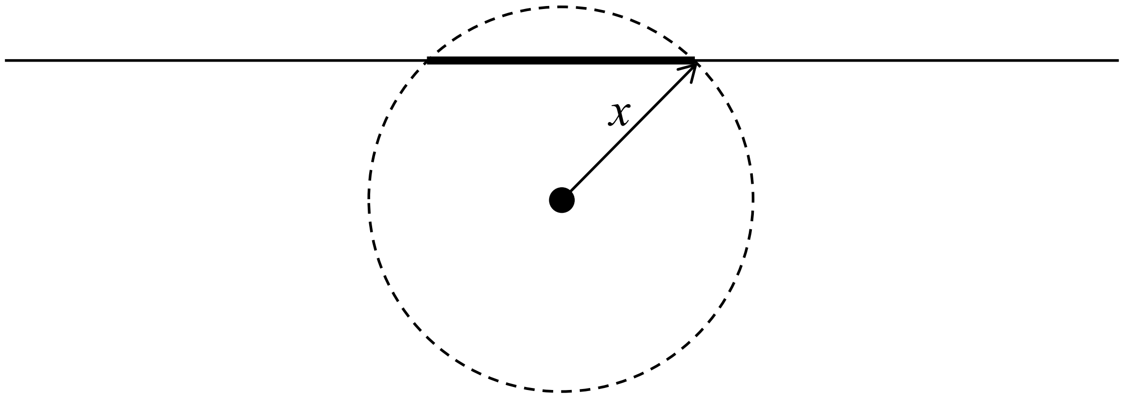
Statistical Inference

Geometrical	Classical	Bayesian
Compare lengths, areas, volumes	Averaging observed data	Consider model parameters as random variables

Geometry

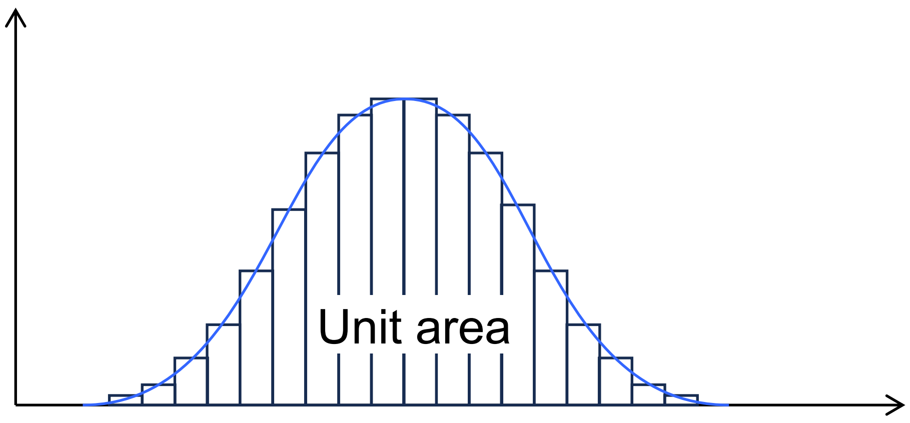


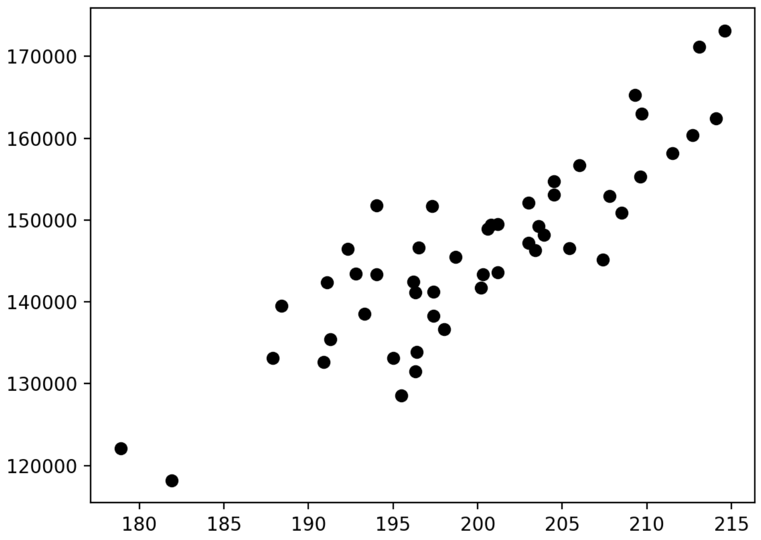
$$F(x) = P(X \le x) = \left\{ \begin{array}{l} \left(\dfrac{L_{X \le x}}{L_{total}} \right) \\ \left(\dfrac{A_{X \le x}}{A_{total}} \right) \\ \left(\dfrac{V_{X \le x}}{V_{total}} \right) \end{array} \right.$$

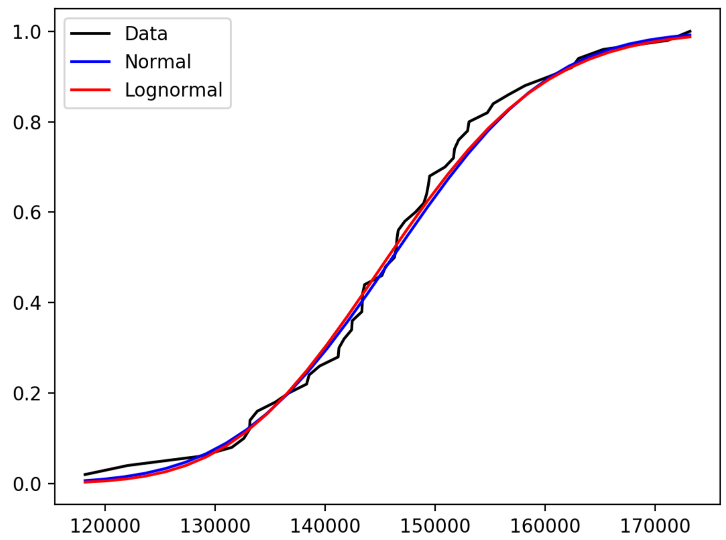


Diagrams

- Histogram
- Frequency diagram
- Cumulative frequency diagram
- Scatter diagram







Mean

Average, expectation

$$\overline{x} = \frac{1}{n} \cdot \sum_{i=1}^{n} x_i$$

Standard Deviation

Basic expression:

$$s^2 = \frac{1}{n-1} \cdot \sum_{i=1}^{n} (x_i - \overline{x})^2$$

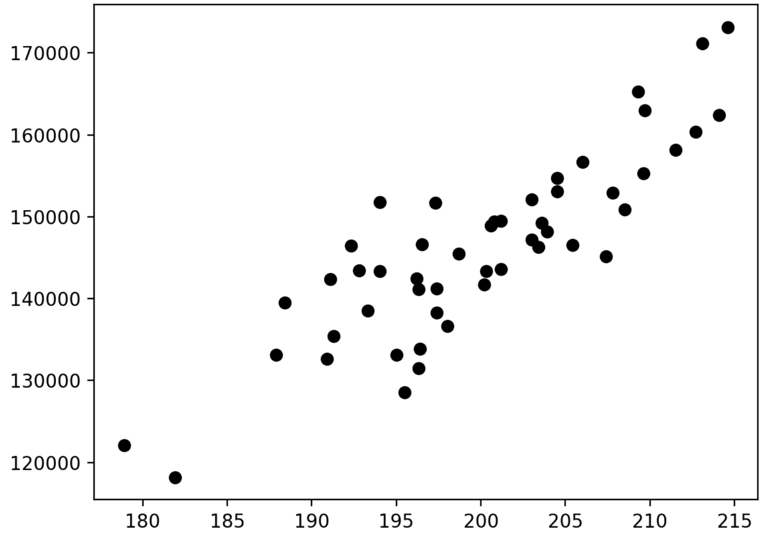
n–1 for the expectation of s to match σ

Computationally more efficient expression:

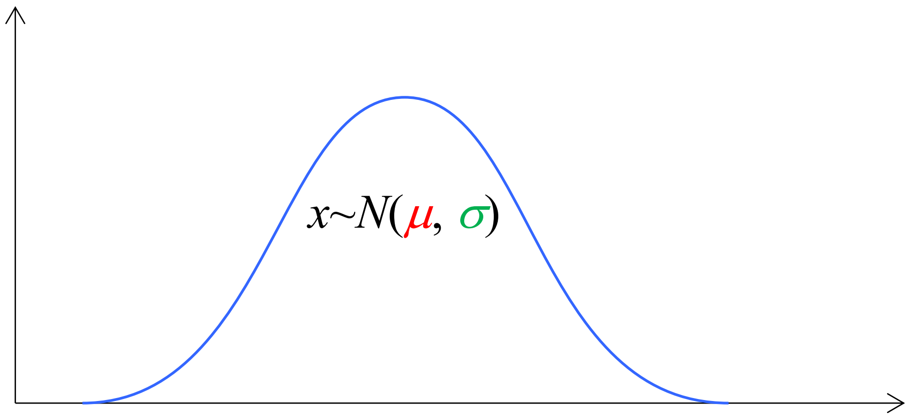
$$\begin{aligned} s^2 &= \frac{1}{n-1} \cdot \sum_{i=1}^{n} (x_i - \overline{x})^2 \\ &= \frac{1}{n-1} \cdot \left(\sum_{i=1}^{n} x_i^2 + \sum_{i=1}^{n} \overline{x}^2 - \sum_{i=1}^{n} 2 x_i \overline{x} \right) \\ &= \frac{1}{n-1} \cdot \left(\left(\sum_{i=1}^{n} x_i^2 \right) + n\overline{x}^2 - 2n\overline{x}^2 \right) \\ &= \frac{1}{n-1} \cdot \left(\left(\sum_{i=1}^{n} x_i^2 \right) - n \cdot \overline{x}^2 \right) \end{aligned}$$

Correlation Coefficient

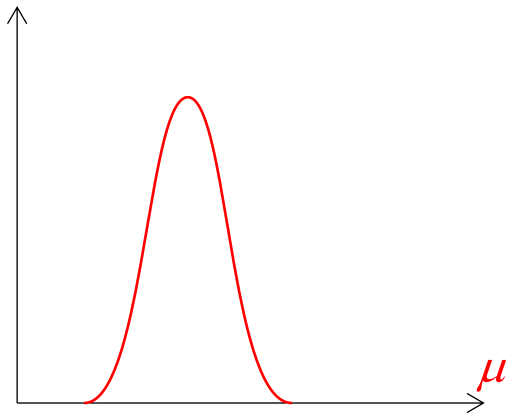
$$\rho = \frac{1}{n-1} \cdot \left(\frac{\sum_{i=1}^{n} x_i y_i - n \cdot \bar{x} \cdot \bar{y}}{s_x \cdot s_y} \right)$$

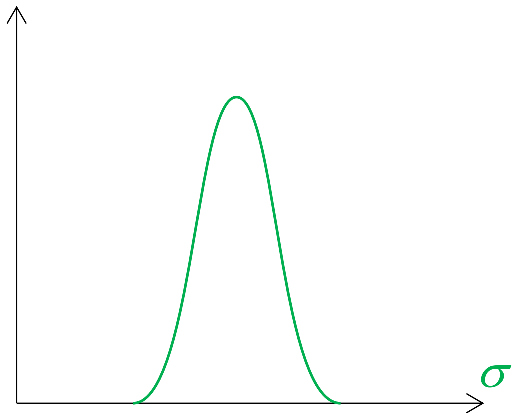


Bayesian Updating



$$P(E_1 \mid E_2) = \frac{P(E_2 \mid E_1)}{P(E_2)} \cdot P(E_1)$$





$$f''(\theta) = \frac{L(\theta)}{c} \cdot f'(\theta)$$

$$c = \int_{-\infty}^{\infty} L(\theta) \cdot f'(\theta) d\theta$$

Likelihood Function

$$f''(\theta) = \frac{L(\theta)}{c} \cdot f'(\theta)$$

"Probability of observing the observation"

Suppose we have observed $X=x$

$L(\theta) \propto \mathrm{P}(X=x \mid \theta) \propto f(x,\theta)$, where θ is, for example, the mean

$\propto p(x,\theta)$, where θ is, for example, the failure probability in Bernoulli trials

Suppose we have observed $X<x$

$L(\theta) \propto \mathrm{P}(X<x \mid \theta) = F(x,\theta)$

Prior

$$f''(\theta) = \frac{L(\theta)}{c} \cdot f'(\theta)$$

Previous posterior

Uniform and non-informative prior

Conjugate prior:

$f''(\theta)$ becomes same type as $f'(\theta)$

Leads to updating rules

Updating Rules for Conjugate Priors

Random variable	Observation	Conjugate prior	Updating rule
X~Binomial(p,n)	x occurrences in n trials	p~Beta(a,b)	$a''=a'+x$ $b''=b'+n-x$
X~Geometric(p)	x trials until first occurrence	p~Beta(a,b)	$a''=a'+1$ $b''=b'+x-1$
X~NegativeBinomial(p,k)	x trials to k^{th} occurrence	p~Beta(a,b)	$a''=a'+k$ $b''=b'+x-k$
X~Poisson(λ,T)	x occurrences in T	λ~Gamma(ν,k)	$k''=k'+x$ $\nu''=\nu'+T$
X~Exp(λ)	n observations of x	λ~Gamma(ν,k)	$k''=k'+n$ $\nu''=\nu'+\Sigma x_i$

More Updating Rules

X~Normal($\underline{\mu}$,$\underline{\sigma}$)	n observations of x	$\underline{\mu}$~Normal(μ,σ)	$\mu'' = \dfrac{\bar{x}\cdot\sigma'^2 + \mu'\left(\dfrac{\underline{\sigma}^2}{n}\right)}{\sigma'^2 + \dfrac{\underline{\sigma}^2}{n}}$ $\sigma'' = \sqrt{\dfrac{\sigma'^2\cdot\left(\dfrac{\underline{\sigma}^2}{n}\right)}{\sigma'^2 + \left(\dfrac{\underline{\sigma}^2}{n}\right)}}$
X~Lognormal(ζ,$\underline{\sigma}$)	n observations of x with average $\bar{x}$	ζ~Normal(μ,σ)	$\mu'' = \dfrac{\overline{\ln(x)}\cdot\sigma'^2 + \mu'\left(\dfrac{\underline{\sigma}^2}{n}\right)}{\sigma'^2 + \dfrac{\underline{\sigma}^2}{n}}$ $\sigma'' = \sqrt{\dfrac{\sigma'^2\cdot\left(\dfrac{\underline{\sigma}^2}{n}\right)}{\sigma'^2 + \left(\dfrac{\underline{\sigma}^2}{n}\right)}}$

Posterior Statistics

$$\mu_\theta = \int_{-\infty}^{\infty} \theta \cdot f''(\theta) d\theta$$

$$\sigma_\theta^2 = \int_{-\infty}^{\infty} (\theta - \mu_\theta)^2 \cdot f''(\theta) d\theta$$

$$\mathbf{M}_\theta = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \boldsymbol{\theta} \cdot f''(\boldsymbol{\theta}) d\boldsymbol{\theta}$$

$$\boldsymbol{\Sigma}_{\theta\theta} = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \left(\boldsymbol{\theta} - \mathbf{M}_\theta\right) \cdot \left(\boldsymbol{\theta} - \mathbf{M}_\theta\right)^T \cdot f''(\boldsymbol{\theta}) d\boldsymbol{\theta}$$

Predictive Distribution

Total probability integration:

$$f(x) = \int_{-\infty}^{\infty} f(x|\theta) \cdot f''(\boldsymbol{\theta}) \, d\boldsymbol{\theta}$$

More lectures:

Terje’s Toobox:

terje.civil.ubc.ca