

Stat 248 – Handout 1 - Elementary Concepts and Methods - Sp 09

Progression:

Scientific question
Experiment
Measurements/data
Probabilities defined?
Statistical model Inferences

Probabilities $P(A), P(B), \dots$

Random variables X, Y, \dots

c.d.f. $F(x) = \text{Prob}(X \leq x)$

Function of r.v. $Z = h(X)$

Expected value

Discrete $\mu_Z = \sum h(x) \text{Prob}(X=x) = E(Z)$

Continuous $\int h(x) f(x) dx$, $f(x)dx = \text{prob}\{x < X < x+dx\}$, dx small

Mean $\mu_X = E(X)$

Variance $\sigma_X^2 = \sigma_{XX} = E(Z)$, $Z = (X - \mu_X)^2$

S.D. σ_X

Covariance $\sigma_{XY} = E(Z)$, $Z = (X - \mu_X)(Y - \mu_Y)$

Correlation $\rho_{XY} = \sigma_{XY} / \sigma_X \sigma_Y$

$-1 \leq \rho_{XY} \leq 1$

$\rho_{XY} = 0$ if X, Y independent $\rho_{XY} = +1$ if $Y = \alpha + \beta X$

Distributions

Discrete

Poisson $P(X=x) = \mu^x \exp(-\mu) / x!$, $x = 0, 1, 2, \dots$

Continuous

Uniform $f(x) = 1/(B-A)$, $A < x < B$

Exponential $\exp(-x)$, $0 < x < \infty$

Normal/Gaussian $\phi(z) = \exp(-z^2/2) / \sqrt{2\pi}$, $-\infty < z < \infty$

Least squares prediction

$$\min_{\alpha, \beta} E(Y - \alpha - \beta X)^2$$

$$\mu_Y = \alpha + \beta \mu_X$$

$$\sigma_{YX} = \beta \sigma_{XX}$$

$$\min = \sigma_{YY} (1 - \rho_{YX}^2)$$

Statistical inference

Sample X_1, \dots, X_n i.i.d.

Estimates

$$\text{c.d.f. } F(x) = \#\{x_i \leq x\}/n$$

$$\mu_X = \bar{x}$$

$$\sigma_{XX} = \sum (x_i - \bar{x})^2 / (n-1) = s_{XX}$$

$$\sigma_{YX} = \sum (y_i - \bar{y})(x_i - \bar{x}) / (n-1) = s_{YX}$$

$$r_{YX} = s_{YX} / (s_X s_Y)$$

Least squares

$$\min \sum (y_i - a - b x_i)^2$$

$$\bar{y} = a + b \bar{x}$$

$$s_{YX} = b s_{XX}$$

Likelihood $L(\theta | x_1, \dots, x_n) = \prod f(x_i | \theta)$

Maximum likelihood estimation

$$\max_{\theta} L(\theta | x_1, \dots, x_n)$$

Margin of error $2 \text{ s.e. } (\hat{\theta})$

Approximate 95% confidence interval $\hat{\theta} \pm \text{ME}$