1 General definitions

1.1 Basic

- Sample variance

$$S_X^2 = \frac{1}{n-1}\sum_1^n (X_i - \bar{X})^2 \tag{1}$$

- Sample correlation coefficient

$$r_{X,Y} = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{(n-1)\sqrt{(S_X^2 S_Y^2)}} \tag{2}$$

- QQ-plot for cumulative distribution function F is the set of points $\left(q_F\left(\frac{i}{n+1}\right), x_{(i)}\right)$, where $q_F(\cdot)$ is the quantile function for the distribution.

- Mean Squared Error (MSE)

$$\mathrm{MSE}(\theta; T(X), g(\theta)) = \mathbf{E}_\theta \left(T(X) - g(\theta)\right)^2 \tag{3}$$

- Bias-variance decomposition

$$\mathrm{MSE}(\theta; T(X)) = \mathrm{var}_\theta T + \left(\mathbf{E}_\theta T(X) - g(\theta)\right)^2 \tag{4}$$

1.2 k-th order statistic $X_{(k)}$

$X_{(k)}$ — $k-th$ order statistic distribution for n i.i.d. variables from continuous distribution F.

$$f_{(k)}(x) = \frac{n!}{(k-1)!(n-k)!}F(x)^{k-1}(1-F(x))^{n-k}f(x) \tag{5}$$

$$F_{(k)}(x) = \sum_{j=k}^n \binom{n}{j} F(x)^j (1-F(x))^{n-j} \tag{6}$$

$$\mathbf{E}F(X_{(k)}) = \frac{k}{n+1} \tag{7}$$

2 Important distributions

- Geometric distribution $G(\theta),\ 0 \leq \theta \leq 1$
  - pdf
$$f_\theta(k) = (1-\theta)^{1-k}\theta \tag{8}$$
  - cdf
$$F_\theta(k) = 1 - (1-\theta)^k \tag{9}$$
- Gamma distribution $\Gamma(\alpha, \lambda),\ \alpha > 0, \lambda > 0$
  - α and λ are known as "shape" and "inverse scale" parameters.
  - pdf
$$p_{\alpha,\lambda}(x) = \frac{x^{\alpha-1}\lambda^\alpha e^{-\lambda x}}{\Gamma(\alpha)} \tag{10}$$
  - cdf (where $\gamma(s,x) = \int_0^x t^{s-1}e^{-t}dt$ — is the "incomplete gamma function")
$$F_{\alpha,\lambda}(x) = \frac{\gamma(\alpha, x\beta)}{\Gamma(\alpha)} \tag{11}$$

3 Estimators

3.1 Maximum of n uniformally distributed statistics

Set up: $X_1, X_2, \ldots X_n$ i.i.d. drown from $U[0,\theta]$, where θ is the parameter of interest.

- $\hat{\theta} = 2\bar{X_n}$
  - method of moments estimator
  - *unbiased*
  - $\mathrm{MSE}(\theta, \hat{\theta}) = \frac{\theta^2}{3n}$, see (5)
- $\frac{n+1}{n}X_{(n)}$ — n-th order statistic, i.e. maximum.
  - $\mathbf{E}_\theta X_{(n)} = \frac{n+1}{n+2}\theta$, see (5)
  - $\mathrm{MSE}(\theta, X_{(n)}) = \frac{2\theta^2}{(n+2)(n+1)}$

- $\frac{n+2}{n+1}X_{(n)}$
  - *unbiased*
  - $\text{MSE}(\theta, \frac{n+2}{n+1}X_{(n)}) = \frac{2\theta^2}{(n+1)^2}$

3.2 Univariate normal distribution

- $(\hat{\mu}, \hat{\sigma^2}) = \left(\bar{X}_n, \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X}_n)\right) = \left(\bar{X}_n, \frac{n-1}{n}S_X^2\right)$
  - maximum likelihood estimator
  - method of moments estimator
  - $\hat{\mu}$ is unbiased
  - $\mathbf{E}_{(\mu,\sigma^2)}\hat{\sigma^2} = \frac{n-1}{n}\sigma^2$