Statistical Decisions and Risks - Reference Handout

Key Formulas & Definitions

Pierre Pudlo

1 Basic Setup

- **Random dataset:** D taking values in data space, with distribution P_θ
- **Model parameter:** θ (unknown)
- **Parameter of interest:** $\beta = \eta(\theta)$ (can differ from θ)
- **Estimator:** $S = s(D)$ with observed value $s_{\text{obs}} = s(d)$

2 Frequentist Approach

2.1 Loss Functions and Risk

Loss function: $L(\beta, s) \geq 0$, minimized when $s = \beta$

Risk function: $R_\theta(S) = \mathbb{E}_\theta[L(S, \beta)]$

Common losses:

- Squared: $L(s, \beta) = (s - \beta)^2 \to$ Risk = MSE
- Absolute: $L(s, \beta) = |s - \beta| \to$ Risk = MAE
- 0-1: $L(s, \beta) = \mathbf{1}\{s \neq \beta\} \to$ Risk = Error probability

2.2 MSE Decomposition

$$\text{MSE} = \text{Bias}^2 + \text{Variance} = \left(\mathbb{E}_\theta[S] - \beta\right)^2 + \text{Var}_\theta[S]$$

2.3 Key Estimators

Binomial Model: $D \sim \mathcal{B}(n, \theta)$, $\beta = \theta$

$$S_{\alpha,\beta} = \frac{D + \alpha}{n + \alpha + \beta}, \quad \alpha, \beta \geq 0$$

$$\text{MSE} = \frac{(\alpha(1-\theta) - \theta\beta)^2 + n\theta(1-\theta)}{(n + \alpha + \beta)^2}$$

Gaussian Model: $D \sim \mathcal{N}(\mu, \sigma^2)^{\otimes n}$, $\beta = \mu$

$$S_{\kappa,m} = \frac{\kappa m + n\bar{D}}{\kappa + n}, \quad \kappa > 0$$

$$\text{MSE} = \frac{\kappa^2(m - \mu)^2 + n\sigma^2}{(\kappa + n)^2}$$

2.4 Optimality Criteria

Admissible: No other estimator S' exists such that $R_\theta(S') \leq R_\theta(S)$ for all θ and $R_\theta(S') < R_\theta(S)$ for some θ.

Minimax: $\sup_\theta R_\theta(S_0) = \inf_s \sup_\theta R_\theta(S)$

3 Bayesian Approach

3.1 Bayes' Theorem

$$\pi(\theta|d) = \frac{p(d|\theta)\pi(\theta)}{p(d)} \quad \text{where } p(d) = \int p(d|\theta)\pi(\theta)d\theta$$

- $\pi(\theta)$: **prior**
- $p(d|\theta)$: **likelihood**
- $\pi(\theta|d)$: **posterior**
- $p(d)$: **evidence**

3.2 Key Conjugate Pairs

Beta-Binomial:

- Prior: $\theta \sim \text{Beta}(\alpha, \beta)$
- Likelihood: $D|\theta \sim \mathcal{B}(n, \theta)$
- Posterior: $\theta|d \sim \text{Beta}(\alpha + d, \beta + n - d)$
- Posterior mean: $\frac{\alpha+d}{\alpha+\beta+n}$

Normal-Normal: (known σ^2)

- Prior: $\mu \sim \mathcal{N}(m, \kappa^{-2})$
- Likelihood: $D|\mu \sim \mathcal{N}(\mu, \sigma^2)^{\otimes n}$
- Posterior: $\mu|d \sim \mathcal{N}\left(\frac{\kappa^2 m + n\bar{d}/\sigma^2}{\kappa^2 + n/\sigma^2}, \frac{1}{\kappa^2 + n/\sigma^2}\right)$

3.3 Precision Parameterization (Gaussian)

Key insight: Precisions add!

- Prior precision: κ^2
- Data precision: n/σ^2
- Posterior precision: $\kappa^2 + n/\sigma^2$

Posterior mean: $\frac{\kappa^2 \cdot m + (n/\sigma^2) \cdot \bar{d}}{\kappa^2 + n/\sigma^2}$

3.4 Bayesian Risk and Bayes Estimators

Bayesian risk: $R_\pi(S) = \int R_\theta(S)\pi(\theta)d\theta$

Bayes estimator: $s(d) = \arg\min_s \mathbb{E}[L(s, \eta(\theta_{rv}))|D = d]$

Optimal estimators by loss:

- **Squared loss:** Bayes estimator = posterior mean
- **Absolute loss:** Bayes estimator = posterior median
- **Limit case:** Bayes estimator = posterior mode

3.5 Concrete Bayes Estimators

Binomial + Beta + squared loss:

$$\text{Bayes estimator} = \frac{\alpha + d}{\alpha + \beta + n}$$

Gaussian + Normal + squared loss:

$$\text{Bayes estimator} = \frac{\kappa^2 m + n\bar{d}/\sigma^2}{\kappa^2 + n/\sigma^2}$$

4 Hypothesis Testing as Decision Theory

4.1 Setup

- Parameter space: $\beta \in \{0, 1, \dots, K-1\}$
- 0-1 loss: $L(s, \beta) = \mathbf{1}\{s \neq \beta\}$
- Posterior probabilities: $p_i(d) = \mathbb{P}(\beta = i | D = d)$

4.2 Bayes Decision Rule

$$s^*(d) = \arg\max_i p_i(d)$$

For $K = 2$: Choose $s = 1$ if $p_1(d) > 1/2$, otherwise $s = 0$

4.3 Connection to Classical Testing

Frequentist risk interpretation:

$$R_\theta(S) = \mathbb{P}_\theta(S \neq \beta)$$

- When $\beta = 0$: Risk = Type I error
- When $\beta = 1$: Risk = 1 - Power

5 Key Connections

5.1 Admissibility Results

- **Bayes estimators** with proper priors are **admissible**
- **Constant risk** often indicates **minimax** property

5.2 Minimax Results

- **Binomial:** $\dfrac{D + \sqrt{n}/2}{n + \sqrt{n}}$ is minimax
- **Gaussian:** Sample mean $\bar{D}$ is minimax (when σ^2 known)

---

6 Quick Reference: Common Distributions

Distribution	Parameters	Mean	Variance
$\mathcal{B}(n, \theta)$	$n \in \mathbb{N}$, $\theta \in [0, 1]$	$n\theta$	$n\theta(1-\theta)$
$\mathcal{N}(\mu, \sigma^2)$	$\mu \in \mathbb{R}$, $\sigma^2 > 0$	μ	σ^2
$\text{Beta}(\alpha, \beta)$	$\alpha, \beta > 0$	$\frac{\alpha}{\alpha+\beta}$	$\frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$

7 Useful Identities

Sample mean: $\bar{D} = \frac{1}{n}\sum_{i=1}^n D_i$

For i.i.d. sample: $\text{Var}[\bar{D}] = \frac{\sigma^2}{n}$

Standard normal CDF: $\Phi(z) = \mathbb{P}(Z \leq z)$ where $Z \sim \mathcal{N}(0, 1)$