Final Exam – Review

1 Basic concept

1. Expected value $E(X)$
$$E(X) = \sum_x xP(X = x) \quad \text{or} \quad E(X) = \int_{-\infty}^{\infty} xf(x)dx$$

2. Variance $\text{Var}(X)$
$$\text{Var}(X) = E(X^2) - [E(X)]^2$$

3. Covariance $\text{Cov}(X, Y)$
$$\text{Cov}(X, Y) = E(XY) - E(X)E(Y)$$

4. Correlation coefficient between X and Y
$$\rho = \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}(X)\text{Var}(Y)}} \in [-1, 1]$$

5. Law of total expectation
$$E(X) = E[E(X|Y)] \quad \text{or} \quad E(Y) = E[E(Y|X)]$$

6. Covariance between $\sum_{i=1}^{n} a_i X_i$ and $\sum_{j=1}^{m} b_j Y_j$
$$\text{Cov}\left(\sum_{i=1}^{n} a_i X_i, \sum_{j=1}^{m} b_j Y_j\right) = \sum_{i=1}^{n}\sum_{j=1}^{m} a_i b_j \text{Cov}(X_i, Y_j)$$

7. Moment generating function $M(t) = E(e^{tX})$ or joint moment generating function

8. Fisher information
$$I(\theta) = -E\left(\frac{\partial^2 \log(f(X;\theta))}{\partial\theta^2}\right) = \text{Var}\left(\frac{\partial \log(f(X;\theta))}{\partial\theta}\right)$$

9. Convergence in distribution $X_n \xrightarrow{D} X$ (two ways)

10. Convergence in Probability $X_n \xrightarrow{P} X$ (two ways)
$$P(|X_n - X| \geq \epsilon) \longrightarrow 0 \quad \text{as} \quad n \longrightarrow \infty$$

2 Estimation

1. MME (Method of moments estimator): three steps

   Example. Let $X_1, X_2, \ldots, X_n$ be an iid sample from the Beta distribution Beta(a, b). Find the MMEs of parameters a and b.

   Solution. STEP 1: Evaluate the first **two** moments of X_1.

   $$\mu_1 = E(X_1) = \int_0^1 x f(x)dx = \int_0^1 x \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} x^{a-1}(1-x)^{b-1}dx = \frac{a}{a+b},$$
   $$\mu_2 = E(X_1^2) = \int_0^1 x^2 f(x)dx = \int_0^1 x^2 \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} x^{a-1}(1-x)^{b-1}dx = \frac{a}{a+b}\frac{a+1}{a+b+1}.$$

   STEP 2: Solve two equations in Step 1 for a and b in terms of μ_1 and μ_2.

   $$a = \frac{\mu_1(\mu_1 - \mu_2)}{\mu_2 - \mu_1^2},$$
   $$b = \frac{(1-\mu_1)(\mu_1 - \mu_2)}{\mu_2 - \mu_1^2}.$$

   STEP 3: Replace μ_1 and μ_2 in Step 2 by $\hat{\mu}_1$ and $\hat{\mu}_2$ respectively to obtain the MMEs of a and b, where $\hat{\mu}_1 = (1/n)\sum_{i=1}^n X_i$ and $\hat{\mu}_2 = (1/n)\sum_{i=1}^n X_i^2$. That is, the MMEs of a and b are given by

   $$\hat{a} = \frac{\hat{\mu}_1(\hat{\mu}_1 - \hat{\mu}_2)}{\hat{\mu}_2 - \hat{\mu}_1^2},$$
   $$\hat{b} = \frac{(1-\hat{\mu}_1)(\hat{\mu}_1 - \hat{\mu}_2)}{\hat{\mu}_2 - \hat{\mu}_1^2}.$$

2. MLE (Maximum likelihood estimator)

   (i) Likelihood function: $L(\boldsymbol{\theta}; \mathbf{X}) = \prod_{i=1}^n f(X_i; \boldsymbol{\theta})$.

   (ii) Log-likelihood function: $l(\boldsymbol{\theta}; \mathbf{X}) = \log[L(\boldsymbol{\theta}; \mathbf{X})]$.

   (iii) Three different cases:

   CASE 1: Solve $\dfrac{\partial l(\boldsymbol{\theta}; \mathbf{X})}{\partial \boldsymbol{\theta}} = \mathbf{0}$.

   CASE 2: The likelihood function is always increasing in θ when θ is in 1-dimensional space.

Case 3: The likelihood function is always decreasing in θ when θ is in 1-dimensional space, like **Problem 7.4.9**.

3. Properties of estimator: Unbiasedness; Variance; Asymptotic property including δ-method; Consistency

4. Efficient estimator (two conditions)

$$E(\hat{\theta}) = \theta \quad \text{and} \quad \text{Var}(\hat{\theta}) = \frac{1}{nI(\theta)}.$$

5. Confidence interval: mean, variance, the difference of two means, ratio of two variances

6. Sufficient statistic & factorization criterion

   (i) $T = T(X_1, X_2, \ldots, X_n)$ is a sufficient statistic for θ **iff**

$$P(X_1 \le x_1, X_2 \le x_2, \ldots, X_n \le x_n \,|\, T = t)$$

   does not depend on θ for all $\theta \in \Omega$

   (ii) $T = T(X_1, X_2, \ldots, X_n)$ is sufficient for θ **iff**

$$\prod_{i=1}^{n} f(x_i; \theta) = k_1\big(t(x_1, x_2, \ldots, x_n); \theta\big)\, k_2(x_1, x_2, \ldots, x_n),$$

   where $k_1 \ge 0$ depends on t and θ and $k_2 \ge 0$ depends on the sample only.

7. Complete family and complete statistic

   (i) We say that the family $\{f(x; \theta) : \theta \in \Omega\}$ is complete if

$$E_\theta[g(X)] = 0 \quad \text{for all} \quad \theta \in \Omega$$

   implies that

$$P_\theta[g(X) = 0] = 1 \quad \text{for all} \quad \theta \in \Omega.$$

   (ii) A statistic $T(\mathbf{X})$ is said to be complete if the family of distributions of $T(\mathbf{X})$ is complete.

8. Exponential family

   (i) We say that the family $\{f(x;\theta) : \theta \in \Omega\}$ is a one-parameter exponential family if there exist real-valued functions $p(\theta)$ and $q(\theta)$ on Ω and functions $K(x)$ and $S(x)$ such that

$$f(x;\theta) = \exp\{p(\theta)K(x) + S(x) + q(\theta)\}.$$

   (ii) We say that the family $\{f(x;\boldsymbol{\theta}) : \boldsymbol{\theta} \in \Omega\}$ is an m-parameter exponential family if there exist real-valued functions $p_1(\boldsymbol{\theta}), \ldots, p_m(\boldsymbol{\theta})$ and $q(\boldsymbol{\theta})$ on Ω and functions $K_1(x), \ldots, K_m(x)$ and $S(x)$ such that

$$f(x;\boldsymbol{\theta}) = \exp\left\{\sum_{j=1}^{m} p_j(\boldsymbol{\theta})K_j(x) + S(x) + q(\boldsymbol{\theta})\right\}.$$

9. UMVUE and method of finding UMVUE

   Three steps to find the UMVUE for the parameter θ (or $\eta = g(\theta)$):

   Step 1. Find the complete and sufficient statistic T for θ.

   Step 2. Find an unbiased estimator h of θ.

   Step 3. Evaluate $E(h|T)$.

10. Linear regression, least square estimators and their properties

Problem 7.4.9 Let $X_1, \ldots, X_n$ be an iid sample from $U(-\theta, 2\theta)$.

(a) The likelihood function of $X_1, \ldots, X_n$ is given by

$$\begin{aligned} f(\mathbf{x};\theta) &= [1/(3\theta)]^n I(-\theta < \min(x_i) \le \max(x_i) \le 2\theta) \\ &= [1/(3\theta)]^n I(\theta > \max\{-\min(x_i), \max(x_i)/2\}, \end{aligned}$$

the MLE $\hat{\theta}$ of θ is $\max\{-\min(x_i), \max(x_i)/2\}$.

(b) Yes. Use (a) and Factorization Theorem.

(c) Yes. One can use the definition of $\hat{\theta}$ to show that $\hat{\theta}$ has a pdf $f(t;\theta) = nt^{n-1}/\theta^n,\ \ t \in [0,\theta]$ which is complete. That is, $\hat{\theta}$ is sufficient for θ and complete. On the other hand, $E[(n+1)\hat{\theta}/n] = [(n+1)/n]E[\hat{\theta}] = \theta$. That is, $(n+1)\hat{\theta}/n$ is also an unbiased estimator of θ. Thus, $(n+1)\hat{\theta}/n$ is the UMVUE of θ.

3 Testing hypothesis

1. Type I error; Type II error, Significance level of the test, Power, Probability of Type II error

2. Rejection region

3. Normal test

4. χ^2-Square test: Variance of a normal sample, Goodness of the fit test; Homogeneity; Independence

5. t-test

6. F-test

7. Likelihood ratio test

$$\Lambda(\mathbf{X}) = \frac{\max_{\boldsymbol{\theta}\in\omega_0} L(\theta;\mathbf{X})}{\max_{\boldsymbol{\theta}\in\Omega} L(\theta;\mathbf{X})}$$

8. Neyman-Pearson Theorem

$$H_0 : \theta = \theta_0 \quad \text{vs} \quad H_1 : \theta = \theta_1$$

We reject H_0 if

$$\frac{L(\theta_0;\mathbf{X})}{L(\theta_1;\mathbf{X})} \le k.$$

9. UMP test

$$H_0 : \theta = \theta_0 \quad \text{vs} \quad H_1 : \theta \in \omega_1 \quad \text{(composite)}$$

Method: Pick a $\theta_1 \in \omega_1$ and apply Neyman-Pearson theorem to perform test for

$$H_0 : \theta = \theta_0 \quad \text{vs} \quad H_1 : \theta = \theta_1.$$

If the rejection region doesn't depend on the choice of θ_1, the test is a UMP test.

4 Quadratic form

1. Distribution of a quadratic form

   Let $\mathbf{Z} \sim N(\mathbf{0}, \mathbf{I}_n)$ and let $\mathbf{A}$ be a symmetric real matrix with rank r. Then

   $$\mathbf{Z}^T\mathbf{A}\mathbf{Z} \sim \chi^2_r \quad \text{if and only if } \ \mathbf{A} \ \text{ is idempotent.}$$

2. Independence of two quadratic forms

   Let $\mathbf{Z} \sim N(\mathbf{0}, \mathbf{I}_n)$ and let $\mathbf{A}$ and $\mathbf{B}$ be two real symmetric matrices. Then $\mathbf{Z}^T\mathbf{A}\mathbf{Z}$ and $\mathbf{Z}^T\mathbf{B}\mathbf{Z}$ are independent if and only if $\mathbf{A}\mathbf{B} = \mathbf{0}$.