Quiz and Answer

1 1st quiz

1.1 Problem 1

Problem statement. Let $X_1, \dots, X_n$ be a sample from the exponential distribution with parameter $\dfrac{2}{\alpha+1}$, i.e. with density

$$f(x;\alpha) = \begin{cases} \dfrac{2}{\alpha+1} \exp\left(-\dfrac{2x}{\alpha+1}\right), & x \geq 0, \\ 0, & x < 0. \end{cases}$$

(a) Find the maximum likelihood estimator $\hat{\alpha}$ of α. Is $\hat{\alpha}$ unbiased and consistent?
(b) Find the method-of-moments estimator $\alpha^{(1)}$ using the first moment. Is $\alpha^{(1)}$ unbiased and consistent?
(c) Find the method-of-moments estimator $\alpha^{(2)}$ using the second moment. Is $\alpha^{(2)}$ unbiased and consistent?
(d) Check the asymptotic normality of $\alpha^{(1)}$ and $\alpha^{(2)}$ and compare these estimators in the mean-square sense.
(e) Compute the Fisher information $I(\alpha)$ and check whether $\alpha^{(1)}$ is R-efficient.

Solution.

(a) MLE, unbiasedness, consistency. The likelihood is

$$L(\alpha) = \prod_{i=1}^{n} f(X_i;\alpha) = \left(\frac{2}{\alpha+1}\right)^n \exp\left(-\frac{2}{\alpha+1}\sum_{i=1}^{n} X_i\right).$$

Let $S = \sum_{i=1}^{n} X_i$ and $\bar{X} = S/n$. The log-likelihood is

$$\ell(\alpha) = n\log 2 - n\log(\alpha+1) - \frac{2S}{\alpha+1}.$$

Differentiate:

$$\ell'(\alpha) = -\frac{n}{\alpha+1} + \frac{2S}{(\alpha+1)^2}.$$

Setting $\ell'(\alpha) = 0$ yields

$$-n(\alpha+1) + 2S = 0 \quad \Longrightarrow \quad \alpha + 1 = \frac{2S}{n} = 2\bar{X},$$

hence

$$\boxed{\hat{\alpha} = 2\bar{X} - 1}.$$

Unbiasedness: for $X \sim \mathrm{Exp}(\lambda)$, $\mathbb{E}[X] = 1/\lambda$. Here $\lambda = 2/(\alpha+1)$, so

$$\mathbb{E}_\alpha[X] = \frac{\alpha+1}{2}, \qquad \mathbb{E}_\alpha[\bar{X}] = \frac{\alpha+1}{2},$$

and therefore

$$\mathbb{E}_\alpha[\hat{\alpha}] = 2\mathbb{E}_\alpha[\bar{X}] - 1 = 2 \cdot \frac{\alpha+1}{2} - 1 = \alpha.$$

So $\hat{\alpha}$ is unbiased.
Consistency: by the law of large numbers, $\bar{X} \xrightarrow{P} (\alpha+1)/2$, hence

$$\hat{\alpha} = 2\bar{X} - 1 \xrightarrow{P} \alpha.$$

So $\hat{\alpha}$ is consistent.

(b) MoM based on the first moment. The theoretical mean is $\mathbb{E}_\alpha[X] = (\alpha+1)/2$. Equating $\bar{X}$ to this mean gives

$$\bar{X} = \frac{\alpha^{(1)}+1}{2} \qquad \Longrightarrow \qquad \boxed{\alpha^{(1)} = 2\bar{X} - 1}.$$

Thus $\alpha^{(1)} = \hat{\alpha}$, hence it is unbiased and consistent as in (a).

(c) MoM based on the second moment. For $X \sim \mathrm{Exp}(\lambda)$, $\mathbb{E}[X^2] = 2/\lambda^2$. With $\lambda = 2/(\alpha+1)$,

$$\mathbb{E}_\alpha[X^2] = \frac{2}{(2/(\alpha+1))^2} = \frac{(\alpha+1)^2}{2}.$$

Let $M_2 = \frac{1}{n}\sum_{i=1}^n X_i^2$. The MoM equation is

$$M_2 = \frac{(\alpha^{(2)}+1)^2}{2} \qquad \Longrightarrow \qquad \alpha^{(2)} + 1 = \sqrt{2M_2} > 0,$$

so

$$\boxed{\alpha^{(2)} = \sqrt{2M_2} - 1 = \sqrt{\frac{2}{n}\sum_{i=1}^n X_i^2} - 1}.$$

Unbiasedness: define $g(t) = \sqrt{2t} - 1$ so that $\alpha^{(2)} = g(M_2)$. Since g is strictly concave on $(0, \infty)$, Jensen's inequality gives

$$\mathbb{E}_\alpha[\alpha^{(2)}] = \mathbb{E}_\alpha[g(M_2)] < g(\mathbb{E}_\alpha[M_2]) = g\left(\frac{(\alpha+1)^2}{2}\right) = \alpha,$$

so $\alpha^{(2)}$ is biased (downward).
Consistency: by LLN, $M_2 \xrightarrow{P} (\alpha+1)^2/2$. Since g is continuous,

$$\alpha^{(2)} = g(M_2) \xrightarrow{P} \alpha.$$

(d) Asymptotic normality and MSE comparison.
Asymptotic normality of $\alpha^{(1)}$: since $\mathrm{Var}_\alpha(X) = (\alpha+1)^2/4$, by CLT,

$$\sqrt{n}(\alpha^{(1)} - \alpha) = 2\sqrt{n}\left(\bar{X} - \frac{\alpha+1}{2}\right) \Rightarrow \mathcal{N}\big(0, (\alpha+1)^2\big),$$

so $\mathrm{AVar}(\alpha^{(1)}) = (\alpha+1)^2/n$.
Asymptotic normality of $\alpha^{(2)}$: using $\mathbb{E}[X^k] = k!/\lambda^k$,

$$\mathrm{Var}_\alpha(X^2) = \frac{24}{\lambda^4} - \left(\frac{2}{\lambda^2}\right)^2 = \frac{20}{\lambda^4} = \frac{5}{4}(\alpha+1)^4.$$

Thus

$$\sqrt{n}\left(M_2 - \frac{(\alpha+1)^2}{2}\right) \Rightarrow \mathcal{N}\left(0, \frac{5}{4}(\alpha+1)^4\right).$$

With $h(t) = \sqrt{2t} - 1$ and $h'((\alpha+1)^2/2) = 1/(\alpha+1)$, the delta method gives

$$\sqrt{n}(\alpha^{(2)} - \alpha) \Rightarrow \mathcal{N}\left(0, \frac{5}{4}(\alpha+1)^2\right),$$

so $\mathrm{AVar}(\alpha^{(2)}) = \frac{5}{4}\frac{(\alpha+1)^2}{n}$.
Hence, asymptotically,

$$\mathrm{AMSE}(\alpha^{(1)}) = \frac{(\alpha+1)^2}{n} < \frac{5}{4}\frac{(\alpha+1)^2}{n} = \mathrm{AMSE}(\alpha^{(2)}).$$

(e) Fisher information and R-efficiency. For one observation,

$$\frac{\partial}{\partial\alpha}\log f(X;\alpha) = -\frac{1}{\alpha+1} + \frac{2X}{(\alpha+1)^2} = \frac{2X-(\alpha+1)}{(\alpha+1)^2}.$$

Since $\mathrm{Var}_\alpha(2X) = (\alpha+1)^2$, we obtain

$$I_1(\alpha) = \mathbb{E}_\alpha\left[\left(\frac{2X-(\alpha+1)}{(\alpha+1)^2}\right)^2\right] = \frac{1}{(\alpha+1)^2}, \qquad I_n(\alpha) = \frac{n}{(\alpha+1)^2}.$$

The CRLB for unbiased estimators is $1/I_n(\alpha) = (\alpha+1)^2/n$. But

$$\mathrm{Var}_\alpha(\alpha^{(1)}) = \mathrm{Var}_\alpha(2\bar{X}-1) = 4\frac{\mathrm{Var}_\alpha(X)}{n} = \frac{(\alpha+1)^2}{n},$$

so $\alpha^{(1)}$ attains the CRLB and is R-efficient.

1.2 Problem 2

Problem statement. Let $X_1, \ldots, X_n$ be a sample from the uniform distribution on $[1, \theta+2]$.
(a) Find the maximum likelihood estimator $\hat{\theta}$ of θ. Is $\hat{\theta}$ unbiased and consistent?
(b) Find the method-of-moments estimator θ^* using the first moment. Is θ^* unbiased and consistent?
(c) Check the asymptotic normality of $\hat{\theta}$ and θ^* and compare these estimators in the mean-square sense.

Solution.

(a) MLE, unbiasedness, consistency. For $X \sim U[1, \theta+2]$,

$$f(x;\theta) = \frac{1}{\theta+1}\mathbf{1}\{1 \le x \le \theta+2\}.$$

The likelihood is

$$L(\theta) = (\theta+1)^{-n}\mathbf{1}\{1 \le X_i \le \theta+2,\ \forall i\}.$$

Let $M = X_{(n)} = \max_i X_i$. The constraint is $\theta \ge M-2$. On $[M-2,\infty)$, $(\theta+1)^{-n}$ is decreasing, hence

$$\boxed{\hat\theta = M-2}.$$

Unbiasedness: set $Y_i = X_i - 1 \sim U[0,b]$ with $b = \theta+1$, so $\hat\theta + 1 = Y_{(n)}$ and

$$\mathbb{E}[Y_{(n)}] = \frac{n}{n+1}b.$$

Thus

$$\mathbb{E}_\theta[\hat\theta] = \frac{n}{n+1}(\theta+1) - 1 = \frac{n\theta-1}{n+1} \ne \theta,$$

so $\hat\theta$ is biased.
Consistency: $Y_{(n)} \xrightarrow{a.s.} b$, hence $\hat\theta \xrightarrow{a.s.} \theta$.

(b) MoM based on the first moment. Since $\mathbb{E}_\theta[X] = (1+\theta+2)/2 = (\theta+3)/2$, equating to $\bar X$ yields

$$\bar X = \frac{\theta^*+3}{2} \quad\Longrightarrow\quad \boxed{\theta^* = 2\bar X - 3}.$$

Unbiasedness:

$$\mathbb{E}_\theta[\theta^*] = 2\mathbb{E}_\theta[\bar X] - 3 = \theta.$$

Consistency: by LLN, $\theta^* \xrightarrow{P} \theta$.

(c) Asymptotic behavior and MSE comparison.
Asymptotic normality of θ^*: since $\mathrm{Var}_\theta(X) = (\theta+1)^2/12$, by CLT,

$$\sqrt{n}(\theta^*-\theta) = 2\sqrt{n}\left(\bar X - \frac{\theta+3}{2}\right) \Rightarrow \mathcal{N}\left(0, \frac{(\theta+1)^2}{3}\right),$$

so $\mathrm{AVar}(\theta^*) = (\theta+1)^2/(3n)$.
Limit law for $\hat\theta$: with $Y_{(n)} = \hat\theta+1$ and $b = \theta+1$,

$$\mathbb{P}\Big(\frac{n}{b}(b - Y_{(n)}) > u\Big) = \Big(1 - \frac{u}{n}\Big)^n \to e^{-u},$$

hence

$$\frac{n}{\theta+1}(\theta - \hat\theta) \Rightarrow \mathrm{Exp}(1),$$

so $\sqrt{n}(\hat\theta-\theta) \Rightarrow 0$ (no non-degenerate $\sqrt{n}$-normal limit).
Exact MSE comparison: for $U[0,b]$,

$$\mathrm{Var}(Y_{(n)}) = \frac{n}{(n+1)^2(n+2)}b^2, \qquad \mathrm{Bias}(\hat\theta) = \mathbb{E}[Y_{(n)} - b] = -\frac{b}{n+1}.$$

Thus

$$\mathrm{MSE}_\theta(\hat\theta) = \frac{n}{(n+1)^2(n+2)}b^2 + \frac{b^2}{(n+1)^2} = \frac{2b^2}{(n+1)(n+2)} = \frac{2(\theta+1)^2}{(n+1)(n+2)}.$$

Since θ^* is unbiased,

$$\mathrm{MSE}_\theta(\theta^*) = \mathrm{Var}(\theta^*) = 4\frac{\mathrm{Var}(X)}{n} = \frac{(\theta+1)^2}{3n}.$$

Therefore $\mathrm{MSE}_\theta(\hat\theta) = O(n^{-2})$ while $\mathrm{MSE}_\theta(\theta^*) = O(n^{-1})$, so $\hat\theta$ is better in MSE sense for large n (indeed for all $n \ge 2$).

1.3 Problem 3

Problem statement. A sample is drawn from the discrete distribution with parameter θ and support $\{-2,-1,1,2\}$:

x	-2	-1	1	2
$p_\theta(x)$	$\frac{1}{6}$	$\frac{1}{3}-\theta$	$\frac{1}{3}+\theta$	$\frac{1}{6}$

The observed counts are $\nu = (4,6,6,4)$ with total sample size $n = 20$.
(a) Estimate θ by maximum likelihood.
(b) Estimate θ by the method of moments (at least two different estimators). Compare these estimators with each other and with the MLE.
(c) Which estimator is more efficient?

Solution.

(a) MLE. Up to factors not depending on θ,

$$L(\theta) \propto \left(\frac{1}{3}-\theta\right)^{\nu_{-1}} \left(\frac{1}{3}+\theta\right)^{\nu_1}.$$

Thus

$$\ell'(\theta) = -\frac{\nu_{-1}}{1/3-\theta} + \frac{\nu_1}{1/3+\theta}.$$

Solving $\ell'(\theta) = 0$ gives

$$\boxed{\hat\theta_{\mathrm{ML}} = \frac{\nu_1 - \nu_{-1}}{3(\nu_1+\nu_{-1})}}.$$

For $\nu_1 = \nu_{-1} = 6$,

$$\boxed{\hat\theta_{\mathrm{ML}} = 0}.$$

(b) Two MoM estimators and comparison.
MoM 1 (mean): compute

$$\mathbb{E}_\theta[X] = (-2)\cdot\frac{1}{6} + (-1)\left(\frac{1}{3}-\theta\right) + 1\left(\frac{1}{3}+\theta\right) + 2\cdot\frac{1}{6} = 2\theta,$$

so

$$\boxed{\tilde\theta_1 = \frac{\bar X}{2}}.$$

For the data, $\bar x = 0$, so $\tilde\theta_1 = 0$.
MoM 2 (cell probability): using $p_\theta(-1) = 1/3-\theta$ and $\hat p_{-1} = \nu_{-1}/n$,

$$\boxed{\tilde\theta_2 = \frac{1}{3} - \frac{\nu_{-1}}{n}}.$$

Similarly from $p_\theta(1) = 1/3+\theta$,

$$\boxed{\tilde\theta_3 = \frac{\nu_1}{n} - \frac{1}{3}}.$$

For the data, $\tilde\theta_2 = 1/30$ and $\tilde\theta_3 = -1/30$.

(c) Efficiency (asymptotic). The score is

$$s(\theta \mid X) = \begin{cases} 0, & X = \pm 2, \\ -\dfrac{1}{1/3 - \theta}, & X = -1, \\ \dfrac{1}{1/3 + \theta}, & X = 1, \end{cases}$$

so

$$I(\theta) = \mathbb{E}_\theta[s(\theta \mid X)^2] = \frac{1}{1/3 - \theta} + \frac{1}{1/3 + \theta} = \frac{6}{1 - 9\theta^2}, \quad I_n(\theta) = \frac{6n}{1 - 9\theta^2}.$$

Hence the CRLB for unbiased estimators is

$$\mathrm{Var}_\theta(T) \geq \frac{1}{I_n(\theta)} = \frac{1 - 9\theta^2}{6n}.$$

The MLE is asymptotically efficient in regular one-parameter models, i.e.

$$\mathrm{AVar}(\hat{\theta}_{\mathrm{ML}}) = \frac{1}{nI(\theta)} = \frac{1 - 9\theta^2}{6n},$$

while the displayed MoM estimators have larger asymptotic variances for interior θ. Therefore $\hat{\theta}_{\mathrm{ML}}$ is the most efficient among the listed estimators in asymptotic MSE sense.

2 2nd quiz

2.1 Problem 1

Problem statement. Let $X_1, \ldots, X_n$ be a sample from $\mathrm{Unif}[0, \theta]$.
(a) Using the statistic $X_{(n)}$, construct an exact confidence interval of significance level α for the parameter θ.
(b) Construct asymptotic confidence intervals for θ at level α, using the statistic $2\bar{X}$.
(c) (1 point)* Same question using the statistic $\overline{X^3}$.

Solution.
(a) Exact CI based on $X_{(n)}$. Let $T := X_{(n)} = \max\{X_1, \ldots, X_n\}$. For $0 \leq t \leq \theta$,

$$\mathbb{P}_\theta(T \leq t) = \mathbb{P}_\theta(X_1 \leq t, \ldots, X_n \leq t) = \left(\frac{t}{\theta}\right)^n.$$

Hence the pivot

$$U := \frac{T}{\theta} \in [0, 1], \qquad \mathbb{P}(U \leq u) = u^n, \quad 0 \leq u \leq 1,$$

has a distribution free of θ.
For an equal-tail $(1 - \alpha)$ exact CI, choose $u_L, u_U \in (0, 1)$ such that

$$\mathbb{P}(U \leq u_L) = \frac{\alpha}{2}, \qquad \mathbb{P}(U \leq u_U) = 1 - \frac{\alpha}{2}.$$

Since $\mathbb{P}(U \leq u) = u^n$, we get

$$u_L = \left(\frac{\alpha}{2}\right)^{1/n}, \qquad u_U = \left(1 - \frac{\alpha}{2}\right)^{1/n}.$$

Then

$$\mathbb{P}\Big(u_L \le \frac{T}{\theta} \le u_U\Big) = 1-\alpha \iff \mathbb{P}\Big(\frac{T}{u_U} \le \theta \le \frac{T}{u_L}\Big) = 1-\alpha.$$

Therefore an exact $(1-\alpha)$ confidence interval is

$$\boxed{\Big[\frac{X_{(n)}}{(1-\alpha/2)^{1/n}}, \frac{X_{(n)}}{(\alpha/2)^{1/n}}\Big]}.$$

(b) Asymptotic CI based on $2\bar{X}$. For $X \sim \mathrm{Unif}[0,\theta]$,

$$\mathbb{E}[X] = \frac{\theta}{2}, \qquad \mathrm{Var}(X) = \frac{\theta^2}{12}.$$

By the CLT,

$$\sqrt{n}\Big(\bar{X} - \frac{\theta}{2}\Big) \Rightarrow \mathcal{N}\Big(0, \frac{\theta^2}{12}\Big),$$

hence

$$\sqrt{n}\big(2\bar{X} - \theta\big) \Rightarrow \mathcal{N}\Big(0, \frac{\theta^2}{3}\Big).$$

Equivalently,

$$\frac{\sqrt{3n}\,(2\bar{X}-\theta)}{\theta} \Rightarrow \mathcal{N}(0,1).$$

Let $z_{1-\alpha/2}$ be the $(1-\alpha/2)$ quantile of $\mathcal{N}(0,1)$. Then asymptotically,

$$\mathbb{P}\left(-z_{1-\alpha/2} \le \frac{\sqrt{3n}\,(2\bar{X}-\theta)}{\theta} \le z_{1-\alpha/2}\right) \approx 1-\alpha,$$

which can be rearranged (for large n so that $1 - z_{1-\alpha/2}/\sqrt{3n} > 0$) to the multiplicative CI

$$\boxed{\Big[\frac{2\bar{X}}{1 + z_{1-\alpha/2}/\sqrt{3n}}, \frac{2\bar{X}}{1 - z_{1-\alpha/2}/\sqrt{3n}}\Big]}.$$

(One may also use the plug-in additive form $2\bar{X} \pm z_{1-\alpha/2}\,(2\bar{X})/\sqrt{3n}$.)

(c) Asymptotic CI based on $\overline{X^3}$. Let $M_3 := \overline{X^3} = \frac{1}{n}\sum_{i=1}^n X_i^3$. For $X \sim \mathrm{Unif}[0,\theta]$ and any $m \ge 0$,

$$\mathbb{E}[X^m] = \frac{\theta^m}{m+1}.$$

Hence

$$\mathbb{E}[X^3] = \frac{\theta^3}{4}, \qquad \mathbb{E}[X^6] = \frac{\theta^6}{7},$$

and

$$\mathrm{Var}(X^3) = \mathbb{E}[X^6] - \big(\mathbb{E}[X^3]\big)^2 = \theta^6\Big(\frac{1}{7} - \frac{1}{16}\Big) = \frac{9}{112}\theta^6.$$

By the CLT,

$$\sqrt{n}\Big(M_3 - \frac{\theta^3}{4}\Big) \Rightarrow \mathcal{N}\Big(0, \frac{9}{112}\theta^6\Big).$$

Define the moment-based estimator $\hat{\theta}_3 := g(M_3)$ with $g(u) = (4u)^{1/3}$ so that $g(\theta^3/4) = \theta$. By the delta method,

$$g'(u) = \frac{4^{1/3}}{3}u^{-2/3}, \qquad g'\Big(\frac{\theta^3}{4}\Big) = \frac{4}{3}\theta^{-2}.$$

Therefore

$$\sqrt{n}\,(\hat{\theta}_3-\theta)\ \Rightarrow\ \mathcal{N}\Big(0,\ (g'(\theta^3/4))^2\cdot \mathrm{Var}(X^3)\Big)=\mathcal{N}\Big(0,\frac{\theta^2}{7}\Big).$$

Equivalently,

$$\frac{\sqrt{7n}\,(\hat{\theta}_3-\theta)}{\theta}\ \Rightarrow\ \mathcal{N}(0,1).$$

Thus an asymptotic $(1-\alpha)$ CI is

$$\boxed{\left[\ \frac{\hat{\theta}_3}{1+z_{1-\alpha/2}/\sqrt{7n}}\ ,\ \frac{\hat{\theta}_3}{1-z_{1-\alpha/2}/\sqrt{7n}}\ \right]\ ,\quad \hat{\theta}_3=(4\overline{X^3})^{1/3}.}$$

2.2 Problem 2

Problem statement. Construct the Bayesian estimate of the parameter θ of a uniform distribution on the interval $[0,\theta]$ if θ is uniformly distributed on $[0,k]$, $k\geq 1$ and volume of sample $n\geq 3$. Prove that the sample mean is not a minimax estimator and find the minimax estimator for $k=1$.

Solution.
Assume squared-error loss $L(\theta,a)=(a-\theta)^2$.
Bayes estimator under $\theta\sim\mathrm{Unif}[0,k]$. Let $T=X_{(n)}=\max_i X_i$. The likelihood is

$$f(x_1,\dots,x_n\mid\theta)=\theta^{-n}\,\mathbf{1}\{0\leq T\leq\theta\}.$$

With prior density $\pi(\theta)=\frac{1}{k}\mathbf{1}\{0\leq\theta\leq k\}$, the posterior density is, for $\theta\in[T,k]$,

$$\pi(\theta\mid x)\propto\theta^{-n}.$$

Normalizing (using $n\geq 2$),

$$\int_T^k\theta^{-n}\,d\theta=\frac{T^{-(n-1)}-k^{-(n-1)}}{n-1}.$$

Thus, for $\theta\in[T,k]$,

$$\pi(\theta\mid x)=\frac{(n-1)\theta^{-n}}{T^{-(n-1)}-k^{-(n-1)}}.$$

The Bayes estimator under squared loss is the posterior mean:

$$\delta_B(x)=\mathbb{E}[\theta\mid x]=\int_T^k\theta\,\pi(\theta\mid x)\,d\theta=\frac{n-1}{T^{-(n-1)}-k^{-(n-1)}}\int_T^k\theta^{-(n-1)}\,d\theta.$$

Since $n\geq 3$,

$$\int_T^k\theta^{-(n-1)}\,d\theta=\frac{T^{-(n-2)}-k^{-(n-2)}}{n-2}.$$

Hence

$$\boxed{\delta_B(x)=\frac{n-1}{n-2}\cdot\frac{T^{-(n-2)}-k^{-(n-2)}}{T^{-(n-1)}-k^{-(n-1)}}\ ,\qquad T=X_{(n)}.}$$

For $k=1$, this becomes (write $T=t\in(0,1]$)

$$\delta_B(x)=\frac{n-1}{n-2}\cdot\frac{t^{-(n-2)}-1}{t^{-(n-1)}-1}=\frac{n-1}{n-2}\cdot t\cdot\frac{1-t^{n-2}}{1-t^{n-1}}.$$

Sample mean is not minimax (for $k = 1$). For $X \sim \mathrm{Unif}[0, \theta]$,

$$\mathbb{E}[\bar{X}] = \frac{\theta}{2}, \qquad \mathrm{Var}(\bar{X}) = \frac{\theta^2}{12n}.$$

Thus the risk of $\delta_{\text{mean}} = \bar{X}$ is

$$R(\theta, \bar{X}) = \mathbb{E}_\theta[(\bar{X} - \theta)^2] = \mathrm{Var}(\bar{X}) + \big(\mathbb{E}[\bar{X}] - \theta\big)^2 = \frac{\theta^2}{12n} + \Big(\frac{\theta}{2} - \theta\Big)^2 = \theta^2\Big(\frac{1}{4} + \frac{1}{12n}\Big).$$

On $\theta \in [0, 1]$, the maximum occurs at $\theta = 1$:

$$\sup_{\theta \in [0,1]} R(\theta, \bar{X}) = \frac{1}{4} + \frac{1}{12n}.$$

Now consider the estimator $\delta_c = c\, X_{(n)}$ for a constant $c > 0$. Since $X_{(n)}/\theta \sim \mathrm{Beta}(n, 1)$,

$$\mathbb{E}[X_{(n)}] = \frac{n}{n+1}\theta, \qquad \mathbb{E}[X_{(n)}^2] = \frac{n}{n+2}\theta^2.$$

Hence

$$R(\theta, cX_{(n)}) = \mathbb{E}_\theta[(cX_{(n)} - \theta)^2] = c^2\mathbb{E}[X_{(n)}^2] - 2c\theta\mathbb{E}[X_{(n)}] + \theta^2 = \theta^2\left(c^2\frac{n}{n+2} - 2c\frac{n}{n+1} + 1\right).$$

On $\theta \in [0, 1]$, $\sup_\theta R(\theta, cX_{(n)})$ is attained at $\theta = 1$ and equals the bracketed expression. Minimizing this quadratic in c gives

$$\frac{d}{dc}\left(c^2\frac{n}{n+2} - 2c\frac{n}{n+1} + 1\right) = 0 \implies c^* = \frac{n+2}{n+1}.$$

The corresponding minimized maximum risk is

$$\sup_{\theta \in [0,1]} R(\theta, c^* X_{(n)}) = 1 - \frac{\left(\frac{n}{n+1}\right)^2}{\frac{n}{n+2}} = 1 - \frac{n(n+2)}{(n+1)^2} = \frac{1}{(n+1)^2}.$$

For $n \geq 3$,

$$\frac{1}{(n+1)^2} < \frac{1}{4} + \frac{1}{12n},$$

so there exists an estimator ($c^* X_{(n)}$) with strictly smaller maximal risk than $\bar{X}$. Therefore $\bar{X}$ cannot be minimax.

Minimax estimator for $k = 1$. By the above explicit minimization of the maximal risk within the natural class $cX_{(n)}$, the candidate is

$$\boxed{\delta^*(X) = \frac{n+2}{n+1} X_{(n)}\,, \qquad (k = 1).}$$

Its maximal risk over $\theta \in [0, 1]$ equals

$$\boxed{\sup_{\theta \in [0,1]} R(\theta, \delta^*) = \frac{1}{(n+1)^2}.}$$

(Under squared loss, any estimator can be Rao–Blackwellized with respect to the sufficient statistic $X_{(n)}$ to not increase risk pointwise, so it is natural to restrict attention to functions of $X_{(n)}$.)

2.3 Problem 3

Problem statement. Find the efficient estimator of the parameter θ among all unbiased estimators, if the sample comes from $\mathrm{Unif}[0,\theta]$, using the statistics $X_{(n)}$. Explain all conclusions.

Solution.
Let $T = X_{(n)}$. The joint density can be written as

$$f(x_1,\dots,x_n \mid \theta) = \theta^{-n}\,\mathbf{1}\{0 \le T \le \theta\},$$

so T is sufficient for θ.
Moreover, T/θ has density $nu^{n-1}\mathbf{1}\{0 \le u \le 1\}$, which does not depend on θ. This implies completeness of T for this one-parameter family: if $\mathbb{E}_\theta[g(T)] = 0$ for all $\theta > 0$, then $\mathbb{E}[g(\theta U)] = 0$ for all θ, forcing $g(\cdot) = 0$ a.s. (standard completeness argument for scale families with U having support $(0,1)$).
We now seek an unbiased estimator of θ that is a function of T. Since $T/\theta \sim \mathrm{Beta}(n,1)$,

$$\mathbb{E}_\theta[T] = \frac{n}{n+1}\theta.$$

Therefore

$$\delta_{\text{unb}}(X) = \frac{n+1}{n}T = \frac{n+1}{n}X_{(n)}$$

is unbiased:

$$\mathbb{E}_\theta[\delta_{\text{unb}}] = \frac{n+1}{n}\mathbb{E}_\theta[T] = \theta.$$

Because T is complete and sufficient, the Lehmann–Scheffé theorem implies that δ_{unb} is the unique UMVU estimator of θ, i.e. it has minimum variance among all unbiased estimators.
Its variance can be computed from $\mathbb{E}[T^2] = \frac{n}{n+2}\theta^2$:

$$\mathrm{Var}(T) = \frac{n}{n+2}\theta^2 - \left(\frac{n}{n+1}\theta\right)^2 = \frac{n}{(n+2)(n+1)^2}\theta^2,$$

hence

$$\mathrm{Var}\left(\frac{n+1}{n}T\right) = \left(\frac{n+1}{n}\right)^2 \mathrm{Var}(T) = \frac{\theta^2}{n(n+2)}.$$

Thus the (UMVU / "efficient among unbiased") estimator is

$$\boxed{\delta_{\text{UMVU}}(X) = \frac{n+1}{n}X_{(n)}, \qquad \mathrm{Var}(\delta_{\text{UMVU}}) = \frac{\theta^2}{n(n+2)}.}$$

2.4 Problem 4

Problem statement. The table shows the results of a math exam. For each column: the grades are A,...,F and the number of students for each grade.

grade	A	B	C	D	E	F
number	5	10	12	7	7	5

Test the goodness-of-fit at significance level $\alpha = 0.2$ for the following distributions:
(a) Discrete uniform distribution;
(b) the distribution law is defined by the following table:

grade	A	B	C	D	E	F
p_i	$\frac{1}{8} - \theta$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{8} + \theta$	$\frac{1}{8} + \theta$	$\frac{1}{8} - \theta$

For each task, justify whether the null hypothesis should be rejected or not based on the test results you obtain. Clearly explain why your conclusion follows from the value of the test statistic.

Solution.
Let observed counts be

$$(O_A, O_B, O_C, O_D, O_E, O_F) = (5, 10, 12, 7, 7, 5), \qquad n = \sum O_i = 46.$$

Use Pearson's chi-square statistic

$$\chi^2 = \sum_i \frac{(O_i - E_i)^2}{E_i},$$

and reject H_0 at level α when $\chi^2 > \chi^2_{1-\alpha,\nu}$, where ν is the relevant degrees of freedom.
(a) Discrete uniform on $\{A, \ldots, F\}$. Under H_0, $p_i = 1/6$ for all i, so $E_i = n/6 = 46/6$. Compute

$$\chi^2_{\text{uni}} = \sum_{i=1}^{6} \frac{(O_i - 46/6)^2}{46/6} \approx 5.1304.$$

Here no parameters are estimated, so $\nu = 6 - 1 = 5$. The critical value is $\chi^2_{0.8,5} \approx 7.2893$. Since $5.1304 < 7.2893$, we *do not reject* H_0 at level $\alpha = 0.2$.
(b) Model with parameter θ. Under $H_0(\theta)$, probabilities are

$$p_A = p_F = \frac{1}{8} - \theta, \quad p_D = p_E = \frac{1}{8} + \theta, \quad p_B = p_C = \frac{1}{4},$$

with constraint $-1/8 \le \theta \le 1/8$. Estimate θ by maximum likelihood. The multinomial log-likelihood (up to constants) is

$$\ell(\theta) = (O_A + O_F) \log\left(\frac{1}{8} - \theta\right) + (O_D + O_E) \log\left(\frac{1}{8} + \theta\right),$$

where $O_A + O_F = 10$ and $O_D + O_E = 14$. Differentiate and set to zero:

$$\ell'(\theta) = -\frac{10}{\frac{1}{8} - \theta} + \frac{14}{\frac{1}{8} + \theta} = 0 \implies 14\left(\frac{1}{8} - \theta\right) = 10\left(\frac{1}{8} + \theta\right) \implies \hat{\theta} = \frac{1}{48}.$$

Thus expected counts are

$$E_A = E_F = 46\left(\frac{1}{8} - \frac{1}{48}\right) = 46 \cdot \frac{5}{48} \approx 4.7917, \quad E_D = E_E = 46\left(\frac{1}{8} + \frac{1}{48}\right) = 46 \cdot \frac{7}{48} \approx 6.7083,$$

and $E_B = E_C = 46/4 = 11.5$. The chi-square statistic is

$$\chi^2_{\text{fit}} = \sum_{i=1}^{6} \frac{(O_i - E_i)^2}{E_i} \approx 0.2609.$$

One parameter is estimated, so $\nu = 6 - 1 - 1 = 4$. The critical value is $\chi^2_{0.8,4} \approx 5.9886$. Since $0.2609 < 5.9886$, we *do not reject* H_0 at level $\alpha = 0.2$.

2.5 Problem 5

Problem statement. The table shows the results of a math exam.

grade	0–10	11–20	21–30	31–40	41–50	51–60
number of students	10000	20000	24000	14000	14000	10000

Since the exam scale has many possible values, we consider the score distribution as continuous. Unfortunately, the data allow reconstructing the empirical distribution function only at a few points. Test the goodness-of-fit with a discrete uniform distribution at a significance level $\alpha = 0.1$. Can the Kolmogorov–Smirnov test convince us to reject the hypothesis, or do the data not contradict it? Clearly explain why your conclusion follows from the value of the test statistic.

Solution.
Let $n = 10000 + 20000 + 24000 + 14000 + 14000 + 10000 = 92000$. Interpret the null as the continuous uniform distribution on $[0, 60]$:

$$H_0: \quad F_0(x) = \begin{cases} 0, & x < 0, \\ x/60, & 0 \le x \le 60, \\ 1, & x > 60. \end{cases}$$

From the grouped counts we can compute the empirical CDF at the class endpoints:

$$\widehat{F}(10) = \frac{10000}{92000} = \frac{5}{46}, \quad \widehat{F}(20) = \frac{30000}{92000} = \frac{15}{46}, \quad \widehat{F}(30) = \frac{54000}{92000} = \frac{27}{46},$$
$$\widehat{F}(40) = \frac{68000}{92000} = \frac{34}{46}, \quad \widehat{F}(50) = \frac{82000}{92000} = \frac{41}{46}, \quad \widehat{F}(60) = 1.$$

The Kolmogorov–Smirnov statistic is

$$D_n = \sup_x \left|\widehat{F}_n(x) - F_0(x)\right|.$$

Even with incomplete within-bin information, we obtain a *lower bound* on D_n by checking deviations at points we can reconstruct. At $x = 20$,

$$F_0(20) = \frac{1}{3}, \qquad \widehat{F}(10) = \frac{5}{46}.$$

Because all observations in $(11, 20]$ occur after 10, the empirical CDF just before reaching that bin endpoint is at most $\widehat{F}(10)$, so the deviation near 20 is at least

$$\left|F_0(20) - \widehat{F}(10)\right| = \frac{1}{3} - \frac{5}{46} = \frac{31}{138} \approx 0.2246.$$

Hence

$$D_n \ge \frac{31}{138} \approx 0.2246.$$

For the two-sided KS test at level $\alpha = 0.1$, the critical value satisfies

$$D_{n,\alpha} \approx \frac{c_{0.1}}{\sqrt{n}} \quad \text{with} \quad c_{0.1} \approx 1.22,$$

so

$$D_{n,0.1} \approx \frac{1.22}{\sqrt{92000}} \approx 0.0040.$$

Since the lower bound 0.2246 is far larger than 0.0040, we must have $D_n > D_{n,0.1}$. Therefore the KS test *does* convince us to reject H_0 at significance level $\alpha = 0.1$.

2.6 Problem 6

Problem statement. (2 points). Student grades from 3 to 5, obtained in Calculus in 2024 and in Probability Theory in 2025, are given. The table shows the number of students for each possible pair of grades.

2024\2025	3	4	5
3	120	80	50
4	70	100	70
5	40	60	90

(a) Check the null hypothesis of independence for this data at a significance level of 0.05.
(b) Test the null hypothesis of homogeneity at a significance level of 0.3. For each task, justify whether the null hypothesis should be rejected or not based on the test results you obtain. Clearly explain why your conclusion follows from the value of the test statistic.

Solution.
Let O_{ij} be the observed counts. The total sample size is

$$n = \sum_{i,j} O_{ij} = 680.$$

Row totals (2024) are $(250, 240, 190)$ and column totals (2025) are $(230, 240, 210)$.
Under independence (and likewise for the homogeneity test in a 3×3 table), expected counts are

$$E_{ij} = \frac{(\text{row total}_i)(\text{col total}_j)}{n}.$$

Thus

$$E = \begin{pmatrix} \frac{250\cdot230}{680} & \frac{250\cdot240}{680} & \frac{250\cdot210}{680} \\ \frac{240\cdot230}{680} & \frac{240\cdot240}{680} & \frac{240\cdot210}{680} \\ \frac{190\cdot230}{680} & \frac{190\cdot240}{680} & \frac{190\cdot210}{680} \end{pmatrix} \approx \begin{pmatrix} 84.5588 & 88.2353 & 77.2059 \\ 81.1765 & 84.7059 & 74.1176 \\ 64.2647 & 67.0588 & 58.6765 \end{pmatrix}.$$

Pearson's chi-square statistic is

$$\chi^2 = \sum_{i=1}^{3} \sum_{j=1}^{3} \frac{(O_{ij} - E_{ij})^2}{E_{ij}} \approx 56.3653.$$

Degrees of freedom are $(r-1)(c-1) = (3-1)(3-1) = 4$.
(a) Independence at level 0.05**.** Critical value: $\chi^2_{0.95,4} \approx 9.4877$. Since $56.3653 > 9.4877$, we *reject* independence at significance level 0.05.
(b) Homogeneity at level 0.3**.** Critical value: $\chi^2_{0.7,4} \approx 4.8784$. Since $56.3653 > 4.8784$, we *reject* homogeneity at significance level 0.3.