Name: Solutions

1. Suppose that a response variable is generated by randomly selecting an individual ω from a population Ω and obtaining the measurement $x = X(\omega) \in \{1, 2, 3\}$ and the proportions of elements of Π taking these values is given by one of three possibilities, that we label by a, b and c, as provided in the following table.

	$x=1$	$x=2$	$x=3$
a	1/3	1/3	1/3
b	1/6	1/6	2/3
c	1/4	1/4	1/2

Suppose we observe a single value x.

(a) (5 marks) Identify the statistical model, namely, provide the sample space $\mathcal{X}$, parameter space Θ and family of distributions f_θ.

$\mathcal{X} = \{1,2,3\}$, $\Theta = \{a,b,c\}$

$f_a(1) = 1/3$, $f_a(2) = 1/3$, $f_a(3) = 1/3$

$f_b(1) = 1/6$, $f_b(2) = 1/6$, $f_b(3) = 2/3$

$f_c(1) = 1/4$, $f_c(2) = 1/4$, $f_c(3) = 1/2$

(b) (5 marks) Provide the sampling distributions of $\theta_{MLE}(x)$.

$\Theta_{MLE}(1) = a$, $\Theta_{MLE}(2) = a$, $\Theta_{ML}(3) = b$

	$P_\Theta(\Theta_{MLE}=a)$	$P_\Theta(\Theta_{MLE}=b)$	$P_\Theta(\Theta_{MLE}=c)$
$\Theta=a$	2/3	1/3	0
$\Theta=b$	1/3	2/3	0
$\Theta=c$	1/2	1/2	0

2

(c) (5 marks) Determine whether or not θ_{MLE} is a sufficient statistic and if so, whether or not it is minimal sufficient.

We have that $f_\cdot(1) = f_\cdot(2)$ and $f_\cdot(1) \neq c f_\cdot(3)$ for any $c>0$ so we can write $f_\theta(x) = f_\theta(T(x))$ where T satis $T(1)=T(2)=1$ and $T(3)=3$ and by factorization any such T is sufficient. Then any 1-1 function of T is sufficient and so θ_{MLE} is a 1-1 function of T and so is sufficient. Since we can compute $\theta_{MLE}(x)$ from the likelihood this proves θ_{MLE} is minimal sufficient.

(d) (5 marks) Suppose that the quantity of interest is $\Psi(\theta) = \theta$ and a loss function is given by $L(a,\psi) = 1 - I_{\{a\}}(\psi)$ and $L(\theta,\psi) = 2 - I_{\{\theta\}}(\psi)$ otherwise. Determine the risk function for θ_{MLE}.

note, should have been $2(1-I_{\{\theta\}}(\psi))$

$$R(a,\theta_{MLE}) = P_a(\theta_{MLE}=a)L(a,a) + P_a(\theta_{MLE}=b)L(a,b)$$
$$= 2/3 \cdot 0 + 1/3 \cdot 1 = 1/3$$
$$R(b,\theta_{MLE}) = 1/3 \cdot 2 + 2/3 \cdot 0 = 2/3$$
$$R(c,\theta_{MLE}) = \tfrac{1}{2}\cdot 2 + 2 \cdot 1/2 = 2$$

(e) (5 marks) If a uniform prior is placed on Θ, then determine the prior risk of θ_{MLE}.

$$r(\theta_{MLE}) = \frac{1}{3}\cdot\frac{1}{3} + \frac{1}{3}\cdot\frac{2}{3} + \frac{1}{3}\cdot 2 = 1$$

(f) (5 marks) Suppose δ is a decision function such that $\delta(1,\cdot)$ is degenerate at $b, \delta(2,\cdot)$ is uniform on $\{a,b\}$ and $\delta(3,\cdot)$ is degenerate at b. Determine the risk function of δ and which of δ and θ_{MLE} is prefered.

$$R(\theta,\delta) = \int_{\{1,2,3\}} \int_{\{a,b,c\}} L(\theta,A)\,\delta(x,dA)\,P_\theta(dx)$$

$\int_{\{a,b,c\}} L(a,A)\,\delta(1,dA) = 1$, $\int_{\{a,b,c\}} L(a,A)\,\delta(2,dA) = \frac{1}{2}$,

$\int_{\{a,b,c\}} L(a,A)\,\delta(3,dA) = 1$, $\therefore R(a,\delta) = \frac{1}{3} + \frac{1}{3}\frac{1}{2} + \frac{1}{3} = \frac{5}{6}$

Similarly $R(b,\delta) = \frac{1}{6}$, $R(c,\delta) = 2$.

We have $R(a,\theta_{MLE}) < R(a,\delta)$, $R(b,\theta_{MLE}) > R(b,\delta)$ and $R(c,\theta_{MLE}) = R(c,\delta)$. Therefore, we do not prefer θ_{MLE} to δ or conversely.

(g) (5 marks) Does δ in (f) depend on the data only through a minimal sufficient statistic? If not, modify δ so that it does and determine the risk function of this modified estimator.

We can force δ to depend on the minimal sufficient statistic T by replacing δ by

$\delta_T(x,A) = E(\delta(\cdot,A)\,|\,T)(T(x))$. Then

$P(x=1\,|\,T)(a) = 1/2 = P(x=2\,|\,T)(a)$, $P(x=3\,|\,T)(a) = 0$

$P(x=1\,|\,T)(b) = 0 = P(x=2\,|\,T)(b)$, $P(x=3\,|\,T)(b) = 1$

So $\delta_T(1,\{a\}) = \frac{1}{2}\delta(1,\{a\}) + \frac{1}{2}\delta(2,\{a\}) = \frac{1}{2}\cdot 0 + \frac{1}{2}\cdot\frac{1}{2} = \frac{1}{4} = \delta_T(2,\{a\})$

$\delta_T(3,\{a\}) = 0$, $\delta_T(1,\{b\}) = \frac{1}{2}\delta(1,\{b\}) + \frac{1}{2}\delta(2,\{b\}) = \frac{1}{2} + \frac{1}{2}\cdot\frac{1}{2} = 3/4 = \delta_T(2,\{b\})$

and so $\delta_T(1,\{c\}) = \delta_T(2,\{c\}) = 0$. Finally $\delta_T(3,\{a\}) = \delta_T(3,\{c\}) = 0$

and $\delta(3,\{b\}) = 1$. This determines δ_T and we have

$R(\theta,\delta_T) = R(\theta,\delta)$ $\forall\theta$ by a result proved in class.

2. Suppose that $x_1, \ldots, x_n$ is a sample from a $N(\mu, \sigma^2)$ distribution where $\mu \in R^1, \sigma^2 > 0$ are unknown. Suppose we wish to estimate the third quartile $\psi = \Psi(\mu, \sigma^2) = \mu + \sigma z_{0.75}$ and we use squared error loss.

(a) (5 marks) Justify why we need only consider nonrandomized estimators that are functions of $(\bar{x}, s^2)$.

We have that $T(x) = (\bar{x}, s^2)$ is sufficient and so δ_T defined by $\delta_T(x, B) = \int_{\mathbb{R}^n} \delta(x, B)\, P(dz \mid T)(T(x))$ depends on the data only through $T(x)$ and $R(\psi, \delta_T) = R(\psi, \delta)$.

Now let δ be a decision fn that depends on the data only through T and define $d_\delta(x) = \int_{\mathbb{R}} a\, \delta(x, da)$ when this exists and recall that $\int_{\mathbb{R}^1} (\psi - a)^2\, \delta(x, da) = \infty$ otherwise which implies $R(\psi, \delta) = \infty$ whenever $d_\delta(x)$ doesn't exist a.s. μ_0(Lebesgue measure). When $d_\delta(x)$ exists a.s. we can define it arbitrarily at points where it doesn't exist and then $R(\psi, d_\delta) \le R(\theta, \delta)$ by Jensen's inequality. This completes the proof.

(b) (5 marks) Determine an unbiased estimator of ψ. (Hint: what is $E_{(\mu,\sigma^2)}(s)$ equal to?)

When $X \sim$ Chi-squared(λ) then $E(X^{1/2}) = \int_0^\infty \frac{(\frac{1}{2})^{\frac{\lambda}{2}}}{\Gamma(\lambda/2)} x^{\frac{\lambda+1}{2}} e^{-\frac{x}{2}} dx$

$= 2^{1/2} \frac{\Gamma(\frac{\lambda+1}{2})}{\Gamma(\frac{\lambda}{2})} \int_0^\infty \frac{(\frac{1}{2})^{\frac{\lambda+1}{2}}}{\Gamma(\frac{\lambda+1}{2})} x^{\frac{\lambda+1}{2} - 1} e^{-x/2} dx = 2^{1/2}\, \Gamma(\frac{\lambda+1}{2}) / \Gamma(\lambda/2)$.

Therefore, $E(s) = \frac{\sigma}{\sqrt{n-1}} E\left(\left(\frac{(n-1)s^2}{\sigma^2}\right)^{1/2}\right) = \frac{\sigma}{\sqrt{n-1}}\, 2^{1/2} \frac{\Gamma(\frac{n-1+1}{2})}{\Gamma(\frac{n-1}{2})}$

since $\frac{(n-1)s^2}{\sigma^2} \sim$ Chi-squared$(n-1)$. Therefore

$\bar{x} + \frac{\Gamma(\frac{n-1}{2})}{\Gamma(\frac{n}{2})} \sqrt{\frac{n-1}{2}}\, s\, z_{0.75}$ is an unbiased estimator of ψ.

A

(c) (5 marks) Is the estimator determined in (b) optimal? Justify your conclusion.

The estimator determined in (b) is optimal unbiased. This is because $(\bar{x}, s^2)$ is a complete minimal sufficient statistic and the result follows by Lehmann-Scheffé. There are however estimators with smaller risk. For example consider estimators of the form $d_c^*(x) = \bar{x} + cs$ for some constant c. Then $R(\mu, d^*)$

$$= E_{(\mu,\sigma^2)}\left(\bar{x} + cs - \mu - \sigma z_{.75}\right)^2 = Var_{(\mu,\sigma^2)}(\bar{x})$$

$$+ 2\, Cov_{(\mu,\sigma^2)}(\bar{x}, cs + \sigma z_{.75}) + E_{(\mu,\sigma^2)}(cs - \sigma z_{.75})^2$$

0 since $\bar{x}, s^2$ independent

$$= \sigma^2/n + c^2\sigma^2 - 2 c z_{.75}\sigma\left(\frac{\sigma}{\sqrt{n-1}} 2^{1/2}\,\Gamma\left(\tfrac{n}{2}\right)/\Gamma\left(\tfrac{n-1}{2}\right)\right) + \sigma^2 z_{.75}^2$$

$$= c^2\sigma^2 - 2\sigma^2\sqrt{\frac{2}{n-1}}\, z_{.75}\frac{\Gamma\left(\frac{n}{2}\right)}{\Gamma\left(\frac{n-1}{2}\right)}\, c + \left(\frac{1}{n} + z_{.75}^2\right)\sigma^2$$

and this is minimized when

$$c = \sqrt{\frac{2}{n-1}}\,\frac{\Gamma\left(\frac{n}{2}\right)}{\Gamma\left(\frac{n-1}{2}\right)} z_{.75} \neq \frac{\Gamma\left(\frac{n-1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)}\sqrt{\frac{n-1}{2}}\, z_{.75}$$

so the optimal unbiased estimator is not optimal.

③ (a) $E|X-a_i| = \int_{\{X > a_i\}} (X(\omega)-a_i)P(d\omega) - \int_{\{X < a_i\}} (X(\omega)-a_i)P(d\omega)$

Hence when $m \le a_1 \le a_2$ we have $E|X-a_2| - E|X-a_1|$

$= \int_{\{X>a_2\}} X(\omega)P(d\omega) - a_2 P(X>a_2) - \int_{\{X<a_2\}} X(\omega)P(d\omega) + a_2 P(X<a_2)$

$- \int_{\{X>a_1\}} X(\omega)P(d\omega) + a_1 P(X>a_1) + \int_{\{X<a_1\}} X(\omega)P(d\omega) - a_1 P(X<a_1)$

$= -\int_{\{a_1<X\le a_2\}} X(\omega)P(d\omega) - \int_{\{a_1 \le X<a_2\}} X(\omega)P(d\omega) + (a_1-a_2)P(X>a_1)$

$+ a_2 P(a_1 < X \le a_2) + (a_2-a_1)P(X<a_1) + a_2 P(a_1 \le X < a_2)$

$= 2\left[a_2 P(a_1 < X < a_2) - \int_{\{a_1<X<a_2\}} X(\omega)P(d\omega)\right] + (a_2-a_1)\left[P(X \le a_1) - P(X>a_1)\right]$

$\ge 2\left[a_2 P(a_1 < X < a_2) - a_2 P(a_1 < X < a_2)\right] + (a_2-a_1)\left[P(X \le a_1) - P(X>a_1)\right]$

$\ge 2(a_2-a_1)\left[P(X\le a_1) - .5\right] \ge 0$ since $P(X \le a_1) \ge .5$

Similarly the inequality holds when $a_2 \le a_1 \le m$.

(b) $R(\theta, d) = E_\theta\left[|\beta(\theta) - d(s)|\right]$

$\le E_\theta\left[|\beta(\theta') - d(s)|\right]$ since by (a) the right-hand side of the inequality is minimized when $\beta(\theta')$ is a median of the distribution of d.

Also $|a - (\alpha x + (1-\alpha)y)| = |\alpha(a-x) + (1-\alpha)(a-y)|$

$\le \alpha|a-x| + (1-\alpha)|a-y|$ for $\alpha \in [0,1]$ by the triangle inequality and so the loss function is convex.

As such the decision problem is convex and we can restrict to nonrandomized d.

4. (a) We have that $(\theta_1 \cdot \theta_2)\cdot\theta_3 = \theta_1\theta_2\theta_3 = \theta_1\cdot(\theta_2\cdot\theta_3)$ and $e = 1$ satisfies $1\cdot\theta_1 = \theta_1 \cdot 1 = \theta_1$, while $\theta^{-1} = 1/\theta$ satisfies $\theta\cdot(1/\theta) = (1/\theta)\cdot\theta = 1$. Therefore Θ is a group

(b) Now if $x \in R^n$ then $T_\theta(x/\theta) = x$ and so $T_\theta : R^n \xrightarrow{\text{onto}} R^n$. Also if $T_\theta x_1 = T_\theta x_2$ then $\theta x_1 = \theta x_2$ and so $x_1 = x_2$ and T_θ is 1-1. Finally $T_{\theta_1}(T_{\theta_2})x = T_{\theta_1}\theta_2 x = \theta_1\theta_2 x = T_{\theta_1\theta_2}x$. Therefore this defines an action of Θ on R^n.

If $x \sim f_\theta$ then the density of x is $\theta^{-n} f(x/\theta)$. Now put $y = T_{\theta'}x = \theta' x$ and $J_{\theta'}(x) = (\theta')^{-n}$ so the density of y is $(\theta'\theta)^{-n} f(y/\theta'\theta)$ which is in the model. Therefore Θ is a a symmetry group of the model.

Clearly $\overline{T}_g\,\theta = g\theta$ (from the above) and this gives the action on Θ, namely, just group product).

Fix $\theta_0 \in \Theta$. Then it is clear that $G\theta_0 = \Theta$ and the group acts transitively on Θ.

(c) Under what conditions on x is it true that $T_{\theta_1}x = T_{\theta_2}x$ implies $\theta_1 = \theta_2$. Clearly if $\exists\, x_i \neq 0$ in x, then $\theta_1 = \theta_2$. So we need to delete the point 0 from R^n

The orbit of x is then the ray $Gx = \{\theta x : \theta > 0\}$ in the direction of x.

(d) Clearly $\bar{T}_g^* \theta = g\theta$ since $\Theta = \mathcal{A}$. Also $L(\bar{T}_g\theta, \bar{T}_g^* t) = L(g\theta, gt)$ $= \rho(gt/g\theta) = \rho(t/\theta) = L(\theta, t)$ and so the group leaves the decision problem invariant.

For an estimator $d : \mathbb{R}^n \to \mathcal{A}$ to be equivariant it must satisfy $d(T_g \underset{\sim}{x}) = d(g\underset{\sim}{x}) = T_g^* d(\underset{\sim}{x}) = g\, d(\underset{\sim}{x})$.

(e) We have that $[T_g x] = [g\underset{\sim}{x}] = \|g\underset{\sim}{x}\| = g\|\underset{\sim}{x}\|$ and so $[\,]$ is a transformation variable. Accordingly a maximal invariant is given by $D(x) = T^{-1}_{[x]}\underset{\sim}{x} = \underset{\sim}{x}/\|\underset{\sim}{x}\|$.

(f) We have that $\underset{\sim}{x} = s\underset{\sim}{u}$

$$J_{(s,\underset{\sim}{u})}(\underset{\sim}{x}) = \left| \det \begin{pmatrix} u_1 & s & 0 & \cdots & 0 \\ \vdots & & \ddots & & \\ u_n & & & & s \end{pmatrix} \right|$$

$$= s^{n-1} h(\underset{\sim}{u}) \quad \text{where} \quad h(\underset{\sim}{u}) = |\det(\underset{\sim}{u}\ \underset{\sim}{e}_1 \cdots \underset{\sim}{e}_{n-1})|$$

Therefore the joint density of $s, \underset{\sim}{u}$ is $\theta^{-n} s^{n-1} f(s\underset{\sim}{u}/\theta) h(\underset{\sim}{u})$ and the conditional density of $s \mid \underset{\sim}{u}$ is proportional to $\theta^{-n} s^{n-1} f(s\underset{\sim}{u}/\theta)$

(g) The Pitman estimator d of θ is given by $d(\underset{\sim}{x}) = s\,d(\underset{\sim}{u})$ where $d(\underset{\sim}{u})$ minimizes, taking $\theta = 1$,

$$\int_0^\infty L(1, s\,d(\underset{\sim}{u}))\, s^{n-1} f(s\underset{\sim}{u})\, ds$$

$$= \int_0^\infty \rho(s\,d(\underset{\sim}{u}))\, s^{n-1} f(s\underset{\sim}{u})\, ds.$$

(h) When f is the $N_n(\underset{\sim}{0}, I)$ density then $s^{n-1} f(sy) \propto s^{n-1} \exp\{-\frac{s^2}{2} y'y\}$ $= s^{n-1} \exp\{-s^2/2\}$ since $y'y = 1$. Therefore, using $\rho(\tau) = (\tau - 1)^2$, the integral in (g) equals

$$\int_0^\infty (s d(u) - 1)^2 \, s^{n-1} \exp\{-s^2/2\} \, ds = \int_0^\infty (s^{-1} - d(u))^2 s^{n+1} e^{-\frac{s^2}{2}} ds =$$

~~is minimized by taking d(u) =~~ Therefore, the value of $d(u)$ minimizing the integral is the mean of the distribution of s^{-1} density of $s \propto s^{n+1} e^{-\frac{s^2}{2}}$

$$d(u) = \int_0^\infty s^n e^{-s^2/2} ds \Big/ \int_0^\infty s^{n+1} e^{-\frac{s^2}{2}} ds$$

$u = s^2/2$

$$= 2^{\frac{n-1}{2}} \int_0^\infty u^{\frac{n-1}{2}} e^{-u} du \Big/ 2^{\frac{n}{2}} \int_0^\infty u^{\frac{n}{2}} e^{-u} du.$$

$$= 2^{-\frac{1}{2}} \Gamma\left(\frac{n+1}{2}\right) \Big/ \Gamma\left(\frac{n+2}{2}\right) = c$$

Therefore the Pitman estimate of θ is

$$d(x) = \frac{2^{-\frac{1}{2}} \Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{n+2}{2}\right)} \|x\|.$$

(i) There is no UMPI size α test under G because H_0 is not invariant under G.

(j) We have that $\underset{\sim}{x} = (x_1, \ldots, x_n) \sim \theta^{-n} f(\underset{\sim}{x}/\theta)$

$$= \theta^{-n} (2\pi)^{-n/2} \exp\{-\frac{1}{2\theta^2} \sum_{i=1}^{n} x_i^2\}$$

and so this is equivalent to a 1-parameter exponential family with natural parameter $t = 1/\theta^2$ and $T(\underset{\sim}{x}) = \sum_{i=1}^{n} x_i^2 \sim \theta^2$ chi-squared(n). Also $H_0 : \theta = \theta_0$ vs $H_a : \theta \neq \theta_0$ is equivalent to $H_0 : 1/\theta^2 = 1/\theta_0^2$ vs $H_a : 1/\theta^2 \neq 1/\theta_0^2$. So by the theorem on UMPU tests for such families the test takes the form

$$\varphi(\underset{\sim}{x}) = \begin{cases} 1 & T(\underset{\sim}{x}) \notin (c_1, c_2) \\ \gamma_1 & T(\underset{\sim}{x}) = c_1 \\ \gamma_2 & T(\underset{\sim}{x}) = c_2 \\ 0 & \text{otherwise} \end{cases}$$

and since T has a continuous distribution it takes the form

$$\varphi(\underset{\sim}{x}) = \begin{cases} 1 & T(\underset{\sim}{x}) \notin (c_1, c_2) \\ 0 & \text{otherwise.} \end{cases}$$

where c_1, c_2 satisfy (i) $E_{\theta_0}(\varphi) = 1 - P_{\theta_0}(c_1 < T(x) < c_2) = \alpha$

and (ii) $E_{\theta_0}(T\varphi) = E_{\theta_0}(T) - E_{\theta_0}(I_{(c_1, c_2)}(T)\, T) = \alpha E_{\theta_0}(T)$

Since $E_{\theta_0}(T) = n\theta_0^2$ (ii) is equivalent to

$\int_{c_1}^{c_2} \frac{t}{\theta_0^2} g_n(t/\theta_0^2)\, dt = n(1-\alpha)\theta_0^2$ where g_n is the chi-squared(n) density with cdf G_n. So (i) becomes $G_n(c_2/\sigma_0^2) - G_n(c_1/\sigma_0^2) = 1-\alpha$ and for (ii) $\frac{1}{\sigma_0^2} g_n(\frac{t}{\sigma_0^2}) = (\frac{1}{2\sigma_0^2})^{n/2} \Gamma^{-1}(n/2)\, t^{n/2-1} e^{-t/2\sigma_0^2}$ so

$(t/\sigma_0^2) g_n(t/\sigma_0^2) = (\frac{1}{2\sigma_0^2})^{n/2} \Gamma^{-1}(n/2)\, t^{\frac{n+2}{2}-1} e^{-t/2\sigma_0^2}$

$= (2\sigma_0^2)\, \Gamma(\frac{n+2}{2}) \Gamma^{-1}(\frac{n}{2}) \frac{1}{\sigma_0^2} g_{n+2}(t/\sigma_0^2)$ so (ii) becomes

$(2\sigma_0^2)\Gamma(\frac{n+2}{2})\Gamma^{-1}(\frac{n}{2})\left(G_{n+2}(c_2/\sigma_0^2) - G_{n+2}(c_1/\sigma_0^2)\right) = n(1-\alpha)\sigma_0^2$

or $G_{n+2}(c_2/\sigma_0^2) - G_{n+2}(c_1/\sigma_0^2) = (\frac{n}{2})\Gamma^{-1}(\frac{n+2}{2})\Gamma(\frac{n}{2})(1-\alpha) = 1-\alpha$.

So putting $u_i = c_i/\sigma_0^2$ we need to find (u_1, u_2) solving $\quad G_n(u_2) - G_n(u_1) = G_{n+2}(u_2) - G_{n+2}(u_1) = 1-\alpha$

and this is done numerically.

(b) Note that (u_1, u_2) determined in (a) is the same for every θ_0. So

$$C_\gamma(\underset{\sim}{x}) = \{\theta_0 : \phi_{\theta_0}(\underset{\sim}{x}) = 0\}$$

$$= \{\theta_0 : u_1 \le \frac{1}{\theta_0^2} T(\underset{\sim}{x}) \le u_2\}$$

$$= \{\theta_0 : \frac{T(\underset{\sim}{x})}{u_2} \le \theta_0^2 \le \frac{T(\underset{\sim}{x})}{u_1}\}$$

$$= \left[\sqrt{\frac{T(\underset{\sim}{x})}{u_2}}, \sqrt{\frac{T(\underset{\sim}{x})}{u_1}}\right]$$

is the UMAU size $1-\alpha$ confidence interval for θ_0.