Solutions Homework 6

November 28, 2018

Solution to Exercise 4.1.4: OK, the random observable quantity and the family of possible distributions are already given, call them X and $\{P_\theta : \theta \in \Theta\}$, respectively. The kind of action we want to take is estimate a set, so let the action space consist of $\mathcal{A} = \mathcal{B}_p$, the p-dimensional Borel sets, where p is the dimension of the parameter vector θ. We restrict attention to Borel sets since we know they are Lebesgue measurable, and our next step is to define the loss function $L(\theta, A) = m(A)$, where m is Lebesgue measure on $I\!R^p$. We require that the coverage probability of the confidence set is at least $1 - \alpha$, so we define the space of allowable decision rules

$$\mathcal{D} \;=\; \{\delta : \Xi \longrightarrow \mathcal{A} : P_\theta[\theta \in \delta(X)]\} \geq 1 - \alpha\}, \quad \forall \theta \in \Theta.$$

That seems to be about the sum of this problem.

Solution to Exercise 4.2.5: Exercise 2.3.10 was assigned in Stat 532 last semester, so we will just go through the solution there and pick out what we need for this exercise. There is very little additional work to do. We will use the notations from the solution that was given last semester.

For the Poisson family, the sufficient statistic is $T(x) = x$, and the family is full rank, so this is minimal sufficient.

For the Binomial family, again, $T(x) = x$, and the family is full rank, so again it is minimal sufficient.

For the Beta family, the sufficient statistic is $T(x) = (\log x, \log(1 - x))$. Again, it is full rank, so minimal sufficient.

For the negative binomial, $T(x) = x$ is again minimal sufficient for the same reasons as the previous examples.

Solution to Exercise 4.2.16: Letting $\underline{X}$ and $\underline{Y}$ denote the two independent random vectors corresponding to the two populations, it is clear that we have an exponential family here under all circumstances. So, we can apply the Proposition 4.2.5 to deliver the results on minimal sufficiency.

For the first model where all parameters are unrestricted,

$$\begin{aligned} & f_{\mu_X, \sigma_X^2, \mu_Y, \sigma_Y^2}(\underline{x}, \underline{y}) \\ = \quad & \exp\left[\frac{-1}{2\sigma_X^2}\sum_i x_i^2 + \frac{\mu_X}{\sigma_X^2}\sum_i x_i + \frac{-1}{2\sigma_Y^2}\sum_j y_j^2 + \frac{\mu_Y}{\sigma_Y^2}\sum_j y_j - B(\theta)\right] h(\underline{x}, \underline{y}). \end{aligned}$$

Here, we don't care about the $B(\theta)$ and $h(\underline{x}, \underline{y})$. Our natural parameter and sufficient statistic vectors are given by

$$\begin{aligned} \underline{T} &= \left(\sum_i x_i^2, \sum_i x_i, \sum_j y_j^2, \sum_j y_j\right) \\ \underline{\eta} &= \left(\frac{-1}{2\sigma_X^2}, \frac{\mu_X}{\sigma_X^2}, \frac{-1}{2\sigma_Y^2}, \frac{\mu_Y}{\sigma_Y^2}\right). \end{aligned}$$

Note that the space of values of the natural parameter vector is $(-\infty, 0) \times (-\infty, \infty) \times (-\infty, 0) \times (-\infty, \infty)$, which is a nonempty open set, so has nonempty interior. (It was mentioned in lecture that the Cartesian product of sets with nonempty interior has nonempty interior. Let's verify that. To say a set A has nonempty interior means it has a nonempty open subset, i.e. that there is $x \in A$ such that there is an $\epsilon > 0$ such that for all x_1 such that $\|x_1 - x\| < \epsilon$, we have $x_1 \in A$. The set $\mathcal{N}(x, \epsilon) = \{x_1 : \|x_1 - x\| < \epsilon\}$ is an ϵ-neighborhood of x, or just a neighborhood. So, A having nonempty interior means there is some element with a neighborhood contained in A. Now if two sets A and B each have nonempty interior, then clearly their Cartesian product will have nonempty interior provided the product of two neighborhoods again contains a neighborhood. Hopefully it is clear that the Cartesian product of two subsets is again a subset (of the product set, of course). Also, if we have that $\|(x_1, y_1) - (x, y)\| < \epsilon$, then clearly $\|x_1 - x\| < \epsilon$ and $\|y_1 - y\| < \epsilon$, so $\mathcal{N}((x, y), \epsilon) \subset \mathcal{N}(x, \epsilon) \times \mathcal{N}(y, \epsilon)$.)

Note that we need to verify the identifiability condition, which also implies that T does not satisfy any linear constraints. We know that two normal densities are not the same unless they have the same parameters, and they will be equal at most two points if they have different parameters (set them equal, take logarithms, simplify and the points where they are equal satisfy a polynomial root equation of degree at most 2). Hence, they give different measures, and this follows for products of marginal normal distributions.

In conclusion, for model (i), a minimal sufficient statistic is

$$\underline{T} = \left(\sum_i x_i^2, \sum_i x_i, \sum_j y_j^2, \sum_j y_j \right).$$

Turning to model (ii), assume $\sigma_X^2 = \sigma_Y^2 = \sigma^2$. Then the expontial family representation takes the form

$$\begin{aligned} & f_{\mu_X, \mu_Y, \sigma^2}(\underline{x}, \underline{y}) \\ = \; & \exp\left[\frac{-1}{2\sigma^2} \left\{ \sum_i x_i^2 + \sum_j y_j^2 \right\} + \frac{\mu_X}{\sigma^2} \sum_i x_i + \frac{\mu_Y}{\sigma^2} \sum_j y_j - B(\theta) \right] h(\underline{x}, \underline{y}). \end{aligned}$$

The sufficient statistic and natural parameter vectors are given by

$$\begin{aligned} \underline{T} &= \left(\sum_i x_i^2 + \sum_j y_j^2, \sum_i x_i, \sum_j y_j \right) \\ \underline{\eta} &= \left(\frac{-1}{2\sigma^2}, \frac{\mu_X}{\sigma^2}, \frac{\mu_Y}{\sigma^2} \right). \end{aligned}$$

The space for the natural parameter values is $(-\infty, 0) \times (-\infty, \infty) \times (-\infty, \infty)$, which has nonempty interior, so $\underline{T}$ is minimal sufficient here. One can verify that this model is identifiable similarly to the previous case. (This means $\underline{T}$ does not satisfy any linear constraints, so the family is full rank, and it follows $\underline{T}$ is complete, as well.)

For model (iii), the exponential family has the form

$$\begin{aligned} & f_{\mu,\sigma_X^2,\sigma_Y^2}(\underline{x},\underline{y}) \\ = & \exp\left[\frac{-1}{2\sigma_X^2}\sum_i x_i^2 + \frac{\mu}{\sigma_X^2}\sum_i x_i + \frac{-1}{2\sigma_Y^2}\sum_j y_j^2 + \frac{\mu}{\sigma_Y^2}\sum_j y_j - B(\theta)\right] h(\underline{x},\underline{y}). \end{aligned}$$

natural parameter and sufficient statistic vectors are given by

$$\begin{aligned} \underline{T} &= \left(\sum_i x_i^2, \sum_i x_i, \sum_j y_j^2, \sum_j y_j\right) \\ \underline{\eta} &= \left(\frac{-1}{2\sigma_X^2}, \frac{\mu}{\sigma_X^2}, \frac{-1}{2\sigma_Y^2}, \frac{\mu}{\sigma_Y^2}\right). \end{aligned}$$

The space of values of $\underline{\eta}$ is contained in $(-\infty,0)\times(-\infty,\infty)\times(-\infty,0)\times(-\infty,\infty)$, but of course we cannot expect that the map $\underline{\eta}(\mu,\sigma_X^2,\sigma_Y^2)$ can fill out all of this 4-dimensional set. (It is mathematically possible to map a 3-dimensional set onto a 4-dimensional set with nonempty interior, but you can't write the function as a simple formula.) In particular, we see that

$$\eta_2 = -2\mu\eta_1, \quad \eta_4 = -2\mu\eta_3,$$

so points in $\underline{\eta}(\Theta)$ must satisfy the (nonlinear) constraint

$$\eta_1\eta_4 = \eta_2\eta_3.$$

Note that both sides of this equation are 0 if and only if $\mu = 0$ (since η_1 and η_3 are nonzero, and if both sides are nonzero, then $\mu = -\eta_2/(2\eta_1) = -\eta_4/(2\eta_3)$. Thus, any points $(\eta_1,\eta_2,\eta_3,\eta_4)$ in $(-\infty,0)\times(-\infty,\infty)\times(-\infty,0)\times(-\infty,\infty)$, satisfying the constraint $\eta_1\eta_4 = \eta_2\eta_3$ will in fact be the image of some $\eta(\mu,\sigma_X^2,\sigma_Y^2)$, namely for $\mu = -\eta_2/(2\eta_1)$, $\sigma_X^2 = -1/(2\eta_1)$, and $\sigma_Y^2 = -1/(2\eta_3)$. So, we have the space of values of the natural parameter is

$$\underline{\eta}(\Theta) = \{(\eta_1,\eta_2,\eta_3,\eta_4) : \eta_1 < 0, \eta_3 < 0, \eta_1\eta_4 = \eta_2\eta_3\}.$$

Now, we just have to see if we can find θ_0, …, θ_4 as in the Proposition 2.5. to show that $\underline{T}$ is minimal sufficient. Let's try the following

$$\begin{aligned} \theta_0 &= (0,1/2,1/2) \\ \theta_1 &= (0,1/4,1/2) \\ \theta_2 &= (0,1/2,1/4) \\ \theta_3 &= (1/2,1/2,1/2) \\ \theta_4 &= (1/2,1/4,1/2), \end{aligned}$$

which gives

$$\begin{bmatrix} \underline{\eta}(\theta_0) \\ \underline{\eta}(\theta_1) \\ \underline{\eta}(\theta_2) \\ \underline{\eta}(\theta_3) \\ \underline{\eta}(\theta_4) \end{bmatrix} = \begin{bmatrix} -1 & 0 & -1 & 0 \\ -2 & 0 & -1 & 0 \\ -1 & 0 & -2 & 0 \\ -1 & 1 & -1 & 1 \\ -2 & 2 & -1 & 1 \end{bmatrix}$$

$$\begin{bmatrix} \underline{\eta}(\theta_1) - \underline{\eta}(\theta_0) \\ \underline{\eta}(\theta_2) - \underline{\eta}(\theta_0) \\ \underline{\eta}(\theta_3) - \underline{\eta}(\theta_0) \\ \underline{\eta}(\theta_4) - \underline{\eta}(\theta_0) \end{bmatrix} = \begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 1 \\ -1 & 2 & 0 & 1 \end{bmatrix}.$$

By direct calculation, the determinant of this 4×4 matrix is 1, so we have the desired linear independence. Also, by the same argument as above, the family is identifiable, and hence $\underline{T}$ is minimal sufficient.

Note: $\underline{T}$ is *not* complete in this case since

$$E_\theta \left[n^{-1} T_2 - m^{-1} T_4 \right] \ = \ \mu - \mu \ = \ 0, \quad \forall \theta.$$

Solution to Exercise 4.3.5: The Lebesgue density is

$$f_{\alpha,\beta}(\underline{x}) \ = \ \exp\left[\alpha \sum_i \log x_i \ - \ \frac{1}{\beta} \sum_i x_i \ - \ B(\alpha, \beta) \right] h(\underline{x}).$$

Before going too far, note that β is a scale parameter, and we are asked to show independence of a scale invariant statistic with the sufficient statistic for β. Thus, we will fix α and consider the subfamily

$$f_\beta(\underline{x}) \ = \ \exp\left[-\frac{1}{\beta} \sum_i x_i \ - \ B(\beta) \right] h(\underline{x}),$$

where of course $T = \sum_i X_i$ is complete and sufficient for β by exponential family results. Now

$$V \ = \ \frac{\underline{X}}{T}$$

is scale invariant so its distribution doesn't depend on β. More concretely, if $Z_i = X_i/\beta$, then the Z_i's are i.i.d. $Gamma(\alpha, 1)$, and clearly $V = \underline{Z}/T(\underline{Z})$. Thus, V is ancillary and we conclude that V and T are independent by Basu's theorem.

Solution to Exercise 4.3.7: For part (a), we begin with

$$\begin{aligned} P_\theta[a < X_{(1)} < X_{(n)} \le b] \ &= \ P_\theta[a < X_i \le b, \ 1 \le i \le n] \\ &= \ \prod_{i=1}^n P_\theta[a < X_i \le b] \\ &= \ \left(\frac{b - a}{\theta_2 - \theta_1} \right)^n, \end{aligned}$$

where the last equation holds provided $\theta_1 \le a < b \le \theta_2$. Letting $\underline{T} = (T_1, T_2) = (X_{(1)}, X_{(n)})$, we claim that $\underline{T}$ has a Lebesgue p.d.f. given by

$$\begin{aligned} f_\theta(t_1, t_2) \ &= \ -\frac{\partial^2}{\partial t_1 \partial t_2} \left(\frac{t_2 - t_1}{\theta_2 - \theta_1} \right)^n I_{(\theta_1,\theta_2)\times(\theta_1,\theta_2)}(t_1, t_2) I_{\{(x,y):x<y\}}(t_1, t_2) \\ &= \ n(n-1) \frac{(t_2 - t_1)^{n-2}}{(\theta_2 - \theta_1)^n} I_{(\theta_1,\theta_2)\times(\theta_1,\theta_2)}(t_1, t_2) I_{\{(x,y):x<y\}}(t_1, t_2) \end{aligned}$$

One can check that the proposed density does give the 2-D cdf when appropriately integrated. Now, suppose $h(\underline{T})$ is such that

$$E_\theta[h(\underline{T})] \;=\; 0, \quad \forall \theta.$$

This implies that

$$\int_{\theta_1}^{\theta_2} \int_{\theta_1}^{t_2} h(t_1, t_2)(t_2 - t_1)^{n-2}\, dt_1 dt_2 \;=\; 0, \quad \forall a < b \qquad\qquad (**).$$

From now on, we will replace θ_1 by an a variable (with a possible subscript) and θ_2 by a b variable. Let

$$C \;=\; \{(t_1, t_2) : t_1 < t_2\}.$$

We want to show that

$$m^2\left(\{(t_1, t_2) \in C : t_1 < t_2 \,\&\, h_+(t_1, t_2) \neq h_-(t_1, t_2)\}\right) \;=\; 0. \qquad\qquad (***)$$

If this last displayed result is true, then $h_+ = h_-$ except for a set we don't "care about", and that means $h(T_1, T_2) = 0$, P_θ-a.s. for all θ. To do this, define measures

$$\begin{aligned} \mu_{+,N}(A) &= \int_{A\cap C\cap[-N,N]} h_+(\underline{t})(t_2 - t_1)^{n-2}\, d\underline{t} \\ \mu_{-,N}(A) &= \int_{A\cap C\cap[-N,N]} h_-(\underline{t})(t_2 - t_1)^{n-2}\, d\underline{t} \quad , \end{aligned}$$

where $\underline{t} = (t_1, t_2)$. Our goal is to show that the measures $\mu_{+,N} = \mu_{-,N}$ for all $N > 0$. It then follows from uniqueness in the Radon-Nikodym theorem that (***) holds. We make the restriction to $[-N, N]$ in order to keep each of the measures finite (else we will have a problem with one step in the argument).

From equation (**), we conclude that (***) holds for all A of the form $(a, b) \times (a, b)$ with $a \leq b$. We will show

Claim 1. We can extend this result to show that $\mu_{+,N}(A) = \mu_{-,N}(A)$ holds for all A of the form $(a_1, b_1) \times (a_2, b_2)$ with $a_i \leq b_i$.

Claim 2. The last result extends from A being Cartesian products of intervals to all 2-D Borel sets.

For Claim 1, we need to consider all possible orderings of a_1, b_1, a_2, b_2. Of course, we want $a_i < b_i$, so out of the $4! = 24$ permutations of 4 things, half violate $a_1 < b_1$ and another independent half violate $a_2 < b_2$, so we end up with $24 \times (1/2) \times (1/2) = 6$ cases. Considering each case in turn:

Case 1: $a_2 < b_2 \leq a_1 < b_1$ Since the intersection of $(a_1, b_1) \times (a_2, b_2)$ is empty in this case, both $\mu_{+,N}$ and $\mu_{-,N}$ give such a set measure 0.

Case 2: $a_2 \le a_1 < b_2 \le b_1$ See the upper left panel in the Figure 1. Note that

$$\mu_{+,N}\left((a_1,b_1)\times(a_2,b_2)\right) \;=\; \mu_{+,N}\left((a_1,b_2)\times(a_1,b_2)\right),$$

and the same holds for $\mu_{-,N}$. Here, the rectangle $(a_1,b_2)\times(a_1,b_2)$ is indicated in the figure by the dotted line. Since we know that that $\mu_{+,N}(A)=\mu_{-,N}(A)$ holds for all A of the form $(a,b)\times(a,b)$, we conclude $\mu_{+,N}(A)=\mu_{-,N}(A)$ holds in this case.

Case 3: $a_1 \le a_2 \le b_2 \le b_1$ In this case we have

$$\mu_{+,N}\left((a_1,b_1)\times(a_2,b_2)\right) \;=\; \mu_{+,N}\left((a_1,b_2)\times(a_1,b_2)\right) \;-\; \mu_{+,N}\left((a_1,a_2)\times(a_1,a_2)\right).$$

To see this, look at the upper left subplot in Figure 1. Note that the triangle closed off by the dotted line in the lower left part of the plot is $(a_1,a_2)\times(a_1,a_2)\cap C$, and the larger triangle consisting of this and the shaded area is $(a_1,b_2)\times(a_1,b_2)\cap C$. Of course, the same equality holds for $\mu_{-,N}$, and since $\mu_{+,N}$ and $\mu_{-,N}$ agree for the two sets on the r.h.s. (since these are of the form $(a,b)\times(a,b)$), we conclude they agree in this case.

Case 4: $a_2 \le a_1 \le b_1 \le b_2$ This is similar to the previous case, except that

$$\mu_{+,N}\left((a_1,b_1)\times(a_2,b_2)\right) \;=\; \mu_{+,N}\left((a_2,b_1)\times(a_2,b_1)\right) \;-\; \mu_{+,N}\left((b_1,b_2)\times(b_1,b_2)\right).$$

Case 5: $a_1 \le a_2 \le b_1 \le b_2$ Here we have

$$\begin{aligned}&\mu_{+,N}\left((a_1,b_1)\times(a_2,b_2)\right)\\ &\quad=\; \mu_{+,N}\left((a_2,b_1)\times(a_2,b_1)\right) \;-\; \mu_{+,N}\left((a_1,a_2)\times(a_1,a_2)\right)\\ &\qquad\quad -\,\mu_{+,N}\left((b_1,b_2)\times(b_1,b_2)\right).\end{aligned}$$

Case 6: $a_1 < b_1 \le a_2 < b_2$ In this case

$$\begin{aligned}&\mu_{+,N}\left((a_1,b_1)\times(a_2,b_2)\right)\\ &\quad=\; \mu_{+,N}\left((a_2,b_1)\times(a_2,b_1)\right) \;-\; \mu_{+,N}\left((a_1,a_2)\times(a_1,a_2)\right)\\ &\qquad\quad -\,\mu_{+,N}\left((b_1,b_2)\times(b_1,b_2)\right) \;+\; \mu_{+,N}\left((b_1,a_2)\times(b_1,a_2)\right).\end{aligned}$$

Note that the last term added back in is the small triangle below and to the left of the shaded rectangle in the fourth subplot of Figure 1.

In each of the above cases, we have expressed $\mu_{+,N}(A)$ and $\mu_{-,N}(A)$ for $A=(a_1,b_1)\times(a_2,b_2)$ in terms of sets where $\mu_{+,N}$ and $\mu_{-,N}$ agree. This completes the proof of Claim 1.

Turning to Claim 2, we further claim that if two finite measures ν_1 and ν_2 on $(\Omega,\mathcal{F})$ satisfy $\nu_1(\Omega)=\nu_2(\Omega)$, then

$$\mathcal{G} \;=\; \{A\in\mathcal{F} : \nu_1(A)=\nu_2(A)\}$$

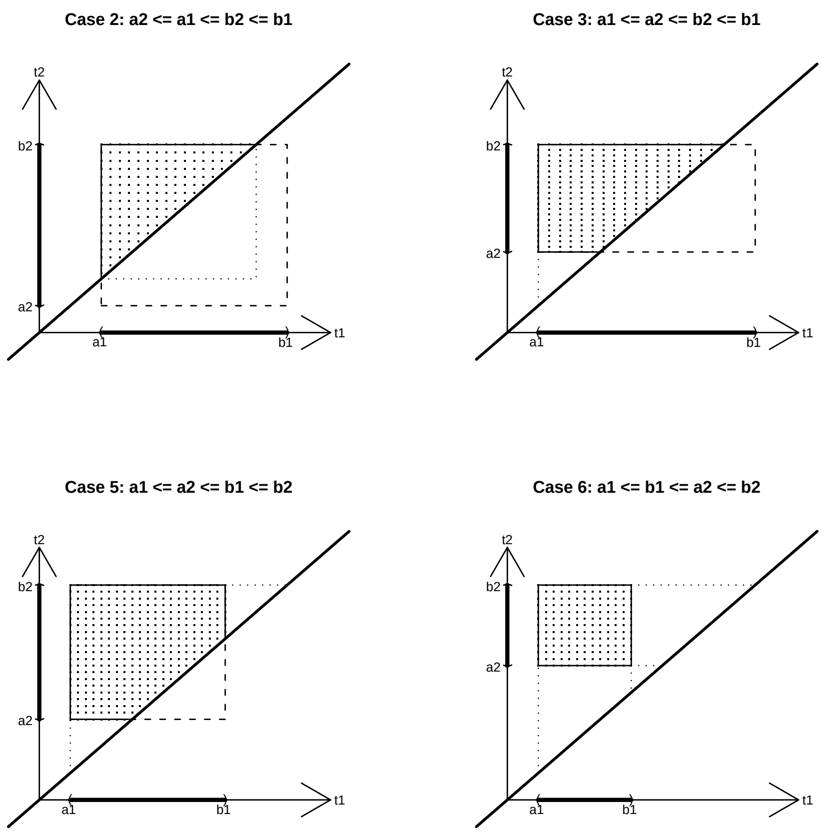

Figure 1: Four of the six cases in verifying Claim 1 of the solution to Exercise 4.3.7.

is a σ-field. Clearly $\emptyset$ and Ω are in $\mathcal{G}$. To show $\mathcal{G}$ is closed under taking complements, let $A \in \mathcal{G}$, then $\nu_1(A^c) = \nu_1(\Omega) - \nu_1(A) = \nu_2(\Omega) - \nu_2(A) = \nu_2(A)$ (here is where we need that the ν_i are finite measures, else we could end up with an $\infty - \infty$ form). Finally, to show that $\mathcal{G}$ is closed under countable unions, it suffices to assume the unions are disjoint, and then apply the countable addivity property.

Thus, we have shown that finite measures $\mu_{+,N}$ and $\mu_{-,N}$ agree on sets of the form $(a_1, b_1) \times (a_2, b_2)$, and they agree on $I\!R^2$ (take $a_1 = a_2 < -N$ and $b_1 = b_2 > N$), so they agree on the σ-field generated by all such rectangles, which is easily shown to be the σ-field of 2-D Borel sets. This completes the proof of Claim 2.

Turning now to part (b) of the exercise, it is easy to check that $\underline{T} = (X_{(1)}, X_{(n)})$ is sufficient (as well as complete): using the factorization theorem,

$$\begin{aligned} f_{\theta_1,\theta_2}(x_1, x_2, \ldots, x_n) &= \prod_{i=1}^{n} (\theta_2 - \theta_1)^{-1} I_{(\theta_1,\theta_2)}(x_i) \\ &= (\theta_2 - \theta_1)^{-n} I_{(\theta_1,\infty)}(\min_i x_i) I_{(-\infty,\theta_2)}(\max_i x_i). \end{aligned}$$

Thus, if we show that

$$\underline{V} = (\underline{X} - X_{(1)})/(X_{(n)} - X_{(1)})$$

is ancillary, then it follows that $\underline{V}$ and $\underline{T}$ are independent by Basu's theorem. Now $\underline{V}$ is clearly location and scale invariant, and our family is a location-scale family (generated by the $Unif(0,1)$ density), so $\underline{V}$ is ancillary. But let's give a more detailed proof of this.

I know that the proofs of ancillarity tend to be short but subtle. Here is one: we can write $\underline{X} \stackrel{D}{=} (\theta_2 - \theta_1)\underline{U} + \theta_1$ where $\underline{U}$ is a vector of U_1, U_2, ..., U_n which are i.i.d. $Unif(0,1)$, and $\stackrel{D}{=}$ means "equal in distribution." Note that the transformation is monotone increasing in each coordinate, so $\underline{T} \stackrel{D}{=} (\theta_2 - \theta_1)(U_{(1)}, U_{(n)}) + \theta_1$. Hence

$$\begin{aligned} \underline{V} &\stackrel{D}{=} \left[(\theta_2 - \theta_1)\underline{U} + \theta_1 - ((\theta_2 - \theta_1)U_{(1)} + \theta_1)\right] / \left[(\theta_2 - \theta_1)U_{(n)} + \theta_1 - ((\theta_2 - \theta_1)U_{(1)} + \theta_1)\right] \\ &= (\underline{U} - U_{(1)})/(U_{(n)} - U_{(1)}), \end{aligned}$$

and the distribution of the last expression does not depend on θ, thus establishing ancillarity of $\underline{V}$.

Solution to Exercise 4.3.8: We want to find a function $h : I\!R^2 \to I\!R$ such that

$$\forall \theta \in I\!R \quad E_\theta \left[h\left(X_{(n)}, X_{(1)}\right)\right] = 0,$$

but yet

$$\exists \theta \in I\!R \quad P_\theta \left[h\left(X_{(n)}, X_{(1)}\right) \neq 0\right] > 0.$$

In fact, we can construct h such that

$$\forall \theta \in I\!R \quad P_\theta \left[h\left(X_{(n)}, X_{(1)}\right) \neq 0\right] = 1.$$

Note that this is a location family generated by $Unif(-1/2, 1/2)$. By switching to the parameter $\theta - 1/2$, we can take the generating distribution to be $Unif(0,1)$, which is slightly easier to work with (then it becomes the Unif$(\theta, \theta + 1)$ family).

Now the sample range $R = X_{(n)} - X_{(1)}$ is a location invariant statistic, so its distribution doesn't depend on the location parameter (i.e., it is ancillary). It is easy to show

$$\begin{aligned} E_\theta\left[X_{(n)}\right] &= \theta + 1 - 1/(n+1), \\ E_\theta\left[X_{(1)}\right] &= \theta + 1/(n+1), \end{aligned}$$

and so

$$\forall \theta \in I\!R \quad E_\theta\left[R - (1 - \frac{2}{n+1})\right] = 0.$$

Thus our function h is

$$h(x,y) = y - x - 1 + \frac{2}{n+1}.$$

Assume for now the reasonable statement that the bivariate r.v. $(X_{(1)}, X_{(n)})$ has a Lebesgue density (in 2 dimensions, of course). Then the event that $(X_{(1)}, X_{(n)})$ satisfies the linear constraint $X_{(n)} - X_{(1)} = 1 - \frac{2}{n+1}$ has probability 0, irrespective of the value of θ. To show that $(X_{(1)}, X_{(n)})$ has a bivariate Lebesgue density, we compute the c.d.f. It suffices to assume the $Unif(0,1)$ distribution since the final result regarding R is invariant to the value of θ. Then the joint c.d.f. is

$$F(x,y) = P[X_{(1)} \le x \,\&\, X_{(n)} \le y].$$

To compute the density we need only consider the values $0 \le x \le y \le 1$. In this region the c.d.f above is equal to

$$\begin{aligned} & P[X_{(n)} \le y] - P[x < X_{(1)} < X_{(n)} \le y] \\ &= P[\text{ all } X_i \le y] - P[\text{ all } X_i \in (x,y)] \\ &= y^n - (y-x)^n. \end{aligned}$$

Applying $\partial^2/\partial x \partial y$ gives the bivariate density by the fundamental theorem of calculus. That theorem tells us integrating this second order partial will give us the c.d.f. we just derived. We don't actually need to compute the density to conclude this.