STAT6201: Theoretical Statistics I

Satvik Saha

Homework 5

1. Let $X \mid \theta \sim f(x \mid \theta)$, $\theta \mid \lambda \sim \pi(\theta \mid \lambda)$, and $\lambda \sim \psi(\lambda)$.

   (a) The mutual information

   $$I(X, Y) = \mathrm{KL}(P_{X,Y}, P_X \otimes P_Y) \geq 0$$

   simply by the properties of the KL divergence (which in turn is due to Jensen's inequality). Furthermore, we have equality if and only if $P_{X,Y} = P_X \otimes P_Y$; but this is precisely the criterion for independence of X and Y.

   (b) Note that $f(x, \lambda) = \psi(\lambda \mid x)m(x)$. We have

   $$\begin{aligned} I(X, \lambda) &= \mathbb{E}_{X,\lambda}\left[\log\left(\frac{f(X,\lambda)}{m(X)\psi(\lambda)}\right)\right] \\ &= \mathbb{E}_X\left[\mathbb{E}_{\lambda|X}\left[\log\left(\frac{f(X,\lambda)}{m(X)\psi(\lambda)}\right)\right]\right] \\ &= \mathbb{E}_X\left[\mathbb{E}_{\lambda|X}\left[\log\left(\frac{\psi(\lambda \mid X)}{\psi(\lambda)}\right)\right]\right] \\ &= \mathbb{E}_X\left[\mathrm{KL}(\psi(\lambda \mid X), \psi(\lambda))\right]. \end{aligned}$$

   (c) The same calculations will give

   $$I(X, \theta) = \mathbb{E}_X\left[\mathrm{KL}(\pi(\theta \mid X), \pi(\theta))\right].$$

   Using the fact that

   $$\mathrm{KL}(\psi(\lambda \mid X), \psi(\lambda)) \leq \mathrm{KL}(\pi(\theta \mid X), \pi(\theta)),$$

   taking expectations with respect to X immediately gives $I(X, \lambda) \leq I(X, \theta)$.

2. Check

   $$\begin{aligned} P(U < \rho(Y)) &= \int P(U < \rho(y))g(y)\,dy \\ &= \int \rho(y)g(y)\,dy \\ &= \frac{1}{M}\int f(y)\,dy \\ &= \frac{1}{M}. \end{aligned}$$

Thus,

$$\begin{aligned}
P(X \in A) &= P(Y \in A \mid U < \rho(Y)) \\
&= \frac{P(Y \in A, U < \rho(Y))}{P(U < \rho(Y))} \\
&= M \int_A \rho(y) g(y)\, dy \\
&= \int_A f(y)\, dy.
\end{aligned}$$

It follows that $X \sim f$.

3. We have the closed unit ball $\mathcal{U}_p$ in $\mathbb{R}^p$, and the unit sphere $\partial\mathcal{U}_p$.

   (a) Define
   $$h \colon \mathbb{R}^p \to \partial\mathcal{U}_p, \qquad x \mapsto x / \|x\|_2\,.$$
   Suppose that Y is spherically symmetric. To show that $h(Y)$ is uniformly distributed on $\partial\mathcal{U}_p$, it suffices to show that $h(Y)$ is spherically symmetric. Indeed, for $H \in O_p(\mathbb{R})$, we have
   $$Hh(Y) = \frac{HY}{\|Y\|_2} = \frac{HY}{\|HY\|_2} = h(HY) \overset{d}{=} h(Y).$$
   The last equality follows since $HY \overset{d}{=} Y$.

   (b) Suppose that $X \sim \text{Uniform}(\mathcal{U}_p)$. Note that its density is
   $$f_X = C\mathbf{1}_{\mathcal{U}_p} = C\mathbf{1}_{[0,1]}(\|\cdot\|_2),$$
   where C is a normalizing constant (the reciprocal of the volume of $\mathcal{U}_p$). Since this is purely a function of $\|x\|_2$, we have X spherically symmetric.

   Next, we claim that $V = h(X)$ and $R = \|X\|_2$ are independent. Recall that V is uniformly distributed on $\partial\mathcal{U}_p$ and is spherically symmetric. Note that for $r \in [0, 1]$, we have
   $$P(R \le r) = P(X \in r\mathcal{U}_p) = \frac{\text{vol}(r\mathcal{U}_p)}{\text{vol}(\mathcal{U}_p)} = r^p.$$
   Thus, we have a density
   $$f_R(r) = pr^{p-1}\,\mathbf{1}_{[0,1]}(r).$$
   Now, for $A \subseteq \partial\mathcal{U}_p$, observe that
   $$P(R \le r, V \in A) = P(X \in \text{cone}_r(A)) = \frac{\text{vol}(\text{cone}_r(A))}{\text{vol}(\mathcal{U}_p)},$$
   where $\text{cone}_r(A) = \bigcup_{r' \in [0,r]} r'A$. But this is just
   $$\frac{\text{vol}(\text{cone}_r(A))}{\text{vol}(\text{cone}_1(A))} \cdot \frac{\text{vol}(\text{cone}_1(A))}{\text{vol}(\mathcal{U}_p)} = r^p \cdot \frac{\text{area}(A)}{\text{area}(\partial\mathcal{U}_p)} = P(R \le r) \cdot P(V \in A).$$
   It follows that V and R are independent.
   Remark: Here, $\text{area}(\cdot)$ refers to the surface area i.e. the Lebesgue measure on $\partial\mathcal{U}_p$. We typically define for measurable $A \subseteq \partial\mathcal{U}_p$
   $$\text{area}(A) = p \cdot \text{vol}(\text{cone}_1(A)).$$

With this, area$(\cdot)$ becomes the measure (up to normalization) describing the distribution of $V = h(X)$, i.e. the uniform distribution on $\partial\mathcal{U}_p$. This can be verified by checking that area$(\cdot)$ is spherically symmetric.

As a consequence, we have

$$\frac{\mathrm{vol}(\mathrm{cone}_1(A))}{\mathrm{vol}(\mathrm{cone}_1(B))} = \frac{\mathrm{area}(A)}{\mathrm{area}(B)}.$$

Remark: With this notation, $\mathrm{cone}_1(\partial\mathcal{U}_p) = \mathcal{U}_p$.

(c) To sample from uniform$(\partial\mathcal{U}_p)$, first generate $Z \sim N(0, \boldsymbol{I}_p)$, say via a vector of $Z_i \overset{iid}{\sim} N(0,1)$. Then, the result in part (a) will show that $V = Z/\|Z\|_2 \sim \mathrm{uniform}(\partial\mathcal{U}_p)$.

Next, we can independently sample $R = U^{1/p}$ where $U \sim \mathrm{uniform}[0,1]$; then, $P(R \leq r) = r^p$, so $R \overset{d}{=} \|X\|_2$ where $X \sim \mathrm{uniform}(\mathcal{U}_p)$. Using the result from (b), we have $VR \overset{d}{=} X \sim \mathrm{uniform}(\mathcal{U}_p)$.

Remark: We have Z spherically symmetric, since it has density

$$f_Z(z) = \frac{1}{(2\pi)^{1/p}} e^{-\|z\|_2^2/2}.$$