Statistics 581, Midterm Exam, Solutions

Wellner; 11/05/2018

1. (24 points) **Define** any **three** of the following six terms.
   (a) The *total variation distance* between two probability measures P and Q.
   (b) The *Hellinger distance* between two probability measures P and Q.
   (c) A normal random vector $Y = (Y_1, \ldots, Y_n)$.
   (d) A *uniformly integrable* sequence of random variables $\{X_n\}$.
   (e) A *standard Brownian motion process* on $[0, 1]$.
   (f) The *inverse or quantile function* F^{-1} of a distribution function F.

   Solution: See Stat 581 course notes, chapters 1 - 3.

2. (30 points). **State** any **three** of the following:
   (a) The Liapunov CLT.
   (b) The Cramér -Wold device.
   (c) The continuous mapping (or Mann-Wald) theorem.
   (d) Vitali's theorem.
   (e) A result connecting the uniform quantile process $\mathbb{V}_n$ to the uniform empirical process $\mathbb{U}_n$.
   (f) Two inequalities relating total variation metric $d_{TV}(P, Q)$ to the Hellinger metric $H(P, Q)$.
   (g) The Helly-Bray theorem.

   Solution: See Stat 581 course notes, chapters 1 - 3.

   Do **either** problem 3 **or** problem 4.

3. (36 points)
   (a) Suppose that $X \sim N(\mu, 1)$. What is the distribution of X^2 when $\mu = 0$? What is the distribution of X^2 when $\mu \neq 0$?
   (b) Suppose that $\underline{X} \sim N_d(\underline{\mu}, I)$ for some vector $\underline{\mu} \in \mathbb{R}^d$. What is the distribution of $Y_d \equiv \sum_{i=1}^d X_j^2 = \|\underline{X}\|^2$?
   (c) For Y_k as in (b), compute $E(Y_k)$ and $Var(Y_k)$.

   Solution: (a) When $X \sim N(\mu, 1)$, $X^2 \sim \chi_1^2(\delta)$, the non-central chi-square distribution with 1 degree of freedom and non-centrality parameter $\delta = \mu^2$. This distribution can also be described conditionally: given $K \sim \text{Poiss}(\delta/2)$, the conditional distribution is $(X^2|K) \sim \chi^2_{1+2K}$.
   (b) When $\underline{X} \sim N_n(\underline{\mu}, I)$, then $Y_n = \|\underline{X}\|^2 \sim \chi_n^2(\delta)$, the non-central chi-square distribution with n degrees of freedom and noncentrality parameter $\delta = \|\underline{\mu}\|^2$. Equivalently, given $K \sim \text{Poiss}(\delta/2)$, the conditional distribution is $(Y_n^2|K) \sim \chi^2_{n+2K}$.
   (c) The mean is $E(Y_n) = E\{E(Y_n|K)\} = E\{n + 2K\} = n + 2(\delta/2) = n + \delta$. The

variance is

$$\begin{aligned} Var(Y_n) &= E\{Var(Y_n|K)\} + Var\{E(Y_n|K)\} \\ &= E\{2(n+2K)\} + Var\{n+2K\} = 2n + 4(\delta/2) + 4(\delta/2) \\ &= 2n + 4\delta. \end{aligned}$$

4. (30 points) Use the Cramér-Chernoff method to find an (exponential) bound for $P(Z \geq z)$ where $Z \sim N(0,1)$.

Solution: For any $r > 0$ we have

$$\begin{aligned} P(Z \geq z) &= P(rZ \geq rz) = P(\exp(rZ) \geq \exp(rz)) \\ &\leq \frac{E \exp(rZ)}{e^{rz}} = \exp(r^2/2 - rz). \end{aligned}$$

This bound holds for all $r > 0$, so we can minimize it with respect to r. Choosing $r = z$ yields $P(Z \geq z) \leq \exp(-z^2/2)$.

5. (36 points)
(a) Define the Hellinger distance $H(P,Q)$ between two probability measures on a common measurable space $(\mathcal{X}, \mathcal{A})$.
(b) Show that $H^2(P,Q) = 1 - \rho(P,Q)$ where $\rho(P,Q) \equiv \int \sqrt{pq}\,d\mu$ for densities p and q of P and Q with respect to some common dominating measure μ.
(c) Now suppose that P is the $N(\mu,1)$ distribution and Q is the $N(\nu,1)$ distribution for some $\mu, \nu \in \mathbb{R}$. Compute $\rho(P,Q)$ in terms of μ and ν.
(d) Suppose that P is the normal $N(0,\sigma^2)$ distribution and Q is the normal $N(0,\tau^2)$ distribution. Show that

$$\rho(P,Q) = \left(\frac{(\sigma^2\tau^2)^{1/2}}{(\sigma^2+\tau^2)/2} \right)^{1/2} \leq 1.$$

Solution: (a) $H^2(P,Q) = (1/2)\int \{\sqrt{p} - \sqrt{q}\}^2 d\mu$ for any measure μ dominating P and Q.
(b) It follows from the definition of H that

$$H^2(P,Q) = (1/2)\int \{p - 2\sqrt{pq} + q\} d\mu = 1 - \int \sqrt{pq}\,d\mu = 1 - \rho(P,Q).$$

(c) When $P = N(\mu,1)$ and $Q = N(\nu,1)$ we have, since

$$\frac{\phi(x-\mu)}{\phi(x)} = \exp(\mu x - (1/2)\mu^2),$$

$$
\begin{aligned}
\rho(P,Q) &= \int \sqrt{\phi(x-\mu)\cdot\phi(x-\nu)}dx = \int \sqrt{\frac{\phi(x-\mu)}{\phi(x)}}\cdot\sqrt{\frac{\phi(x-\nu)}{\phi(x)}}\phi(x)dx \\
&= \int \exp(\mu x/2 - \mu^2/4 + \nu x/2 - \nu^2/4)\phi(x)dx \\
&= \exp\left(-\frac{\mu^2+\nu^2}{4}\right) E\exp\left(\frac{\mu+\nu}{2}Z\right) \\
&= \exp\left(-\frac{\mu^2+\nu^2}{4}\right)\cdot\exp\left(\frac{(\mu+\nu)^2}{4}\cdot\frac{1}{2}\right) \\
&= \exp\left(-\frac{1}{8}(\mu-\nu)^2\right).
\end{aligned}
$$

(d) When P is the $N(0,\sigma^2)$ distribution and Q is the $N(0,\tau^2)$ distribution,

$$
\begin{aligned}
\rho(P,Q) &= \sqrt{\frac{1}{\sigma}\phi(x/\sigma)\cdot\frac{1}{\tau}\phi(x/\tau)dx} \\
&= \frac{1}{\sqrt{\sigma\tau}}\int \phi\left(\frac{x}{\gamma}\right)dx \quad \text{where} \qquad \frac{1}{\gamma^2} = \frac{1}{2}\left(\frac{1}{\sigma^2}+\frac{1}{\tau^2}\right) \\
&= \frac{\gamma}{\sqrt{\sigma\tau}}\int \frac{1}{\gamma}\phi\left(\frac{x}{\gamma}\right)dx \\
&= \frac{\gamma}{\sqrt{\sigma\tau}} \\
&= \left(\frac{(\sigma^2\tau^2)^{1/2}}{(\sigma^2+\tau^2)/2}\right)^{1/2}
\end{aligned}
$$

after some easy algebra. Note that this is the square root of the ratio of the geometric mean of σ^2 and τ^2 to the arithmetic mean, and hence it is ≤ 1 by the Geometric Mean - Arithmetic mean inequality.

Note: A very recent paper, *The total variation distance between high-dimensional Gaussians*, by Luc Devroye, A. Mehrabian, and T. Reddad, arXiv:1810.08693v1, gives bounds on the total variation distance $d_{TV}(N_d(\mu_1,\Sigma_1), N_d(\mu_2,\Sigma_2))$ in terms of $\mu_1, \mu_2, \Sigma_1, \Sigma_2$. Along the way in the proof of their Theorem 1.2 (page 3), they give bounds on the TV distance between $N_1(\mu,\sigma^2)$ and $N_1(\nu,\tau^2)$ when both the means and variances differ.

Do **either** problem 6 **or** problem 7.

6. (36 points).
Suppose that $X, X_1, \ldots, X_n$ are i.i.d. with distribution function F given by $P(X > x) = 1 - F(x) = 1/x^4$, $x \geq 1$, $F(x) = 0$, $x \leq 1$.
(a) For what values of $r > 0$ is $E|X|^r < \infty$? If they are finite compute $\mu = E(X)$ and $\sigma^2 = Var(X)$.
(b) Compute $F^{-1}(t) = Q(t)$, the quantile function corresponding to F.
(c) Which of the following are true? (Briefly indicate why or why not.)

(i) $\sum_{i=1}^n X_i = O_p(n^{1/2})$.

(ii) $\sum_{i=1}^n X_i = O_p(n)$.

(iii) $n^{1/4}(\overline{X}_n - \mu) = o_p(1)$.

(iv) $n^{2/3}(\overline{X}_n - \mu) = O_p(1)$.

(v) $g(n^{1/4}(\overline{X}_n - \mu)) \rightarrow_p 1/2$ where $g(x) = \Phi(x)$, the standard normal distribution function.

(vi) $h(n^{1/2}(\overline{X}_n - \mu)) = O_p(1)$ with $h(x) = 1/|x|$.

$$\sqrt{n}\begin{pmatrix} \mathbb{F}_n^{-1}(1/4) - F^{-1}(1/4) \\ \mathbb{F}_n^{-1}(3/4) - F^{-1}(3/4) \end{pmatrix} \rightarrow_d N_2(0, \Sigma)$$

where

$$\Sigma = \frac{1}{16}\begin{pmatrix} 3Q'(1/4)^2 & Q'(1/4)Q'(3/4) \\ Q'(1/4)Q'(3/4) & 3Q'(3/4)^2 \end{pmatrix}.$$

Solution: (a) We find that

$$\begin{aligned} E|X|^r &= EX^r \quad \text{since} \quad X \geq 1 \quad a.s. \\ &= \int_0^\infty r x^{r-1}(1 - F(x))dx = \int_0^1 r x^{r-1} dx + \int_1^\infty r x^{r-1}(1 - F(x))dx \\ &= 1 + r\int_1^\infty x^{r-1}x^{-4} = 1 + r\int_1^\infty x^{r-5}dx \\ &= 1 + \frac{r}{4 - r} < \infty \end{aligned}$$

if $r < 4$. Taking $r = 1$ yields $\mu = E(X) = 1 + (1/3) = 4/3$, and taking $r = 2$ yields $E(X^2) = 1 + 1 = 2$, so $Var(X) = 2 - (4/3)^2 = (18 - 16)/9 = 2/9$.
(b) The quantile function $Q(u)$ is found by solving $F(Q(u)) = u$, or $1 - F(Q(u)) = 1 - u$, or $1/Q(u)^4 = 1 - u$), and hence $Q(u) = (1 - u)^{-1/4}$.
(c)
(i) is false: since $E(X_1) = 4/3 > 0$, the left side is of order n.
(ii) is true by the WLLN or the SLLN: note that $n^{-1}\sum_{i=1}^n X_i \rightarrow_{p,a.s.} E(X_1) = 4/3$

and hence $n^{-1}\sum_{i=1}^n X_i = O_p(1)$.
(iii) is true: $\sqrt{n}(\overline{X}_n - \mu) \to_d N(0, \sigma^2)$ by the central limit theorem, so $n^{1/4}(\overline{X}_n - \mu) = n^{-1/4}\sqrt{n}(\overline{X}_n - \mu) = o(1)O_p(1) = o_p(1)$.
(iv) is false: $n^{2/3}(\overline{X}_n - \mu) = n^{1/6}\sqrt{n}(\overline{X}_n - \mu) = n^{1/6}O_p(1)$ is unbounded in probability (and a.s. by the Law of the Iterated Logarithm).
(v) is true: $n^{1/4}(\overline{X}_n - \mu) = o_p(1)$ as in (iii). Then $g(n^{1/4}(\overline{X}_n - \mu)) \to_p g(0) = \Phi(0) = 1/2$ by the continuous mapping theorem (since $g = \Phi$ is continuous everywhere.
(vi) is true: since $Y_n n^{1/2}(\overline{X}_n - \mu)) \to_d Y \sim N(0, \sigma^2)$ and h is continuous a.s. P_Y, $h(Y_n) \to_d h(Y) = 1/|Y|$ by the continuous mapping theorem.
(vii) is true: this follows from our theorem about the finite-dimensional distributions of the quantile process upon noting that $Q(u) = (1 - u)^{-1/4}$ is differentiable at $u = 1/4$ and at $u = 3/4$.

7. (36 points; from problem set #4)
Suppose that $X_1, X_2, \ldots$ are i.i.d. positive random variables, and define $\overline{X}_n \equiv n^{-1}\sum_{i=1}^n X_i$, $H_n \equiv 1/(n^{-1}\sum_{i=1}^n (1/X_i))$, and $G_n \equiv \{\prod_{i=1}^n X_i\}^{1/n}$ to be the *arithmetic, harmonic, and geometric* means respectively. We know that $\overline{X}_n \to_{a.s.} E(X_1) = \mu$ if and only if $E|X_1| < \infty$.
(a) Use the SLLN together with appropriate additional hypotheses to show that $H_n \to_{a.s.} 1/\{E(1/X_1)\} \equiv h$, and $G_n \to_{a.s.} \exp(E\{\log X_1\}) \equiv g$.
(b) Use the multivariate CLT and the delta method to find the joint limiting distribution of $\sqrt{n}(\overline{X}_n - \mu, H_n - h, G_n - g)$. You will need to impose or assume additional moment conditions to be able to prove this. Specify these additional assumptions carefully.

Solution: See the solution to HW 4, problem 3.