Statistics 523, Problem Set 1 Solutions

Wellner; 4/9/99

1. Clarify the statement of Theorem 13.3.3, PfS page 294, and then prove it (Exercise 13.3.1).

   Solution: The theorem says: "The random variables $X_1, \ldots, X_k$ are independent if and only if

   $$\phi_{\underline{X}}(\cdot) = \prod_{j=1}^{k} \phi_{X_j}(\cdot)\,.'' \tag{1}$$

   The equality (1) means that

   $$\phi_{\underline{X}}(\underline{t}) = \prod_{j=1}^{k} \phi_{X_j}(t_j) \tag{2}$$

   for all $\underline{t} = (t_1, \ldots, t_k) \in R^k$. Thus the "$\cdot$" means something different on the two sides of the equality (1). To prove the theorem, suppose first that $X_1, \ldots, X_k$ are independent. Then

   $$\begin{aligned} \phi_{\underline{X}}(\underline{t}) &= E\exp(\sum_{j=1}^{k} it_jX_j) = E\prod_{j=1}^{k}\exp(it_jX_j) \\ &= \prod_{j=1}^{k} E\exp(it_jX_j) = \prod_{j=1}^{k}\phi_{X_j}(t_j) \end{aligned}$$

   where we have used Theorem 8.1.1 ($k-1$ times). Now suppose that (2) holds for all $\underline{t} = (t_1, \ldots, t_k) \in R^k$. One way to argue this is simply to note that the right side of (2) is the characteristic function of $(X_1', \ldots, X_k') \sim P_1 \times \cdots \times P_k$ where P_j denotes the probability distribution of X_j on R. Thus (2) says that the characteristic function of $(X_1, \ldots, X_k)$ equals the characteristic function of $(X_1', \ldots, X_k')$ where the X_j''s are independent with the same marginal distributions, $X_j' =_d X_j$. By the uniqueness of characteristic functions for random vectors (claimed on page 294, PfS), this implies that $(X_1, \ldots, X_k) =_d (X_1', \ldots, X_k')$; i.e. $X_1, \ldots, X_k$ are independent. Another way to argue this is to note that the class of functions

$$g(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-isy} \psi(s) ds \tag{3}$$

obtained by letting ψ range over all characteristic functions on R with $\int |\psi(t)| dt < \infty$, is the class of uniformly continuous density functions on R, and more generally for $\psi \in L_1(\lambda)$ for Lebesgue measure λ, then $g \in C_0$. Transforming across (2) coordinatewise using (3) and Fubini's theorem yields

$$E\left\{\prod_{j=1}^{k} g_j(X_j)\right\} = \prod_{j=1}^{k} E g_j(X_j)$$

for $g_j \in C_0(R)$, $j = 1, \ldots, k$. But the class of functions $g(\underline{x}) = \prod_{j=1}^{k} g_j(x_j)$, $g_j \in C_0(R)$ is a determining class of functions for distributions on R^k (note that we can approximate indicator functions of arbitrary bounded rectangles by functions in this class), and it follows that $X_1, \ldots, X_k$ are independent.

2. PfS, Exercise 13.4.1, page 297. Are there comparable bounds for $x > \sqrt{2n-1}$ (with the understanding that $(1 - x/n)^n = [(1 - x/n)^+]^n$. In fact, show that with

$$\Delta_n \equiv \sup_{x \geq 0} |(1 - x/n)^n 1_{[0,n]}(x) - e^{-x}|$$

we have $2e^{-2} \leq n\Delta_n \leq (2 + n^{-1})e^{-2}$ for all $n \geq 1$.

Solution: As mentioned in class on 4/7/99, the lower bound of the first part is false: taking $n = 4$ and $x = 2.6 < 2.646 = \sqrt{7}$ we compute $e^{-x} - (1 - x/n)^n = .05927...$ while $x^2 e^{-x}/(2n) = .628$. (Thanks to Dick Hwang for pointing this out. Several of you apparently found counterexamples.) See the attached plots at the end of this solution set.

Although the claimed inequalities have not yet been proved, it is clear that something like this is probably true because we do have

$$n(e^{-x} - (1 - x/n)^n) \to \frac{1}{2} x^2 e^{-x}$$

for each fixed $x > 0$ as $n \to \infty$. This follows from L'Hopital's rule by writing the limit of the left side as

$$\begin{aligned}\lim_{t\to 0} \frac{e^{-x} - (1-tx)^{1/t}}{t} &= \lim_{t\to 0} \frac{-(1-tx)^{1/t}[-t^{-2}\log(1-xt) - t^{-1}(1-tx)^{-1}x]}{1} \\ &= e^{-x}x^2/2\,.\end{aligned}$$

I do not yet have a proof of the claimed upper bound, but here is one attempt. It suffices to show that

$$1 - e^x\left(1 - \frac{x}{n}\right)^n \le \frac{x^2}{2n-1}$$

if $0 \le x \le \sqrt{2n-1}$, or equivalently that

$$1 - \frac{x^2}{2n-1} \le \left(e^{x/n}\left(1 - \frac{x}{n}\right)\right)^n$$

or, letting $y = x/n \le \sqrt{2n-1}/n$,

$$1 - \frac{1}{2-1/n}ny^2 \le \left(1 - \frac{y^2}{2} - \frac{y^3}{3} - \frac{y^4}{8} - \cdots\right)^n$$

where the coefficient of y^k is $1/[k(k-2)!]$. But since $(1-z)^n \ge 1 - nz$, the right side of the last display is bounded below by

$$1 - n\left(\frac{y^2}{2} + \frac{y^3}{3} + \frac{y^4}{8} + \cdots\right) = 1 - ny^2\sum_{j=0}^{\infty} \frac{y^j}{j!(j+2)} \ge 1 - ny^2\frac{1}{2-1/n}$$

if and only if

$$\sum_{j=0}^{\infty} \frac{y^j}{j!(j+2)} \le \frac{1}{2-1/n}\,.$$

for $y \le \sqrt{2n-1}/n$. Unfortunately, this inequality seems to be incorrect, so my attempted proof of the upper bound fails here. Apparently a bit too much has been given up in one of the earlier inequalities, perhaps $(1-z)^n \ge 1 - nz$.

Now for the second part. Set $f_n(x) \equiv f(x) = e^{-x} - (1 - x/n)^n$ for $0 \le x \le n$. When $n = 1$, $f(x)$ is maximized at $x_1 = 1$ and the maximum value is e^{-1}. Now suppose that $n \geq 2$. Then $f'(x) = -e^{-x} + (1 - x/n)^{n-1} = 0$ if $x = x_n$ satisfies

$$(1 - x_n/n)^{n-1} = e^{-x_n}\,. \tag{4}$$

Then we have

$$\begin{aligned} f(x_n) &= e^{-x_n} - (1 - \frac{x_n}{n})^n \\ &= (1 - \frac{x_n}{n})^{n-1} - (1 - \frac{x_n}{n})^n \\ &= (1 - \frac{x_n}{n})^{n-1} \Big(1 - (1 - \frac{x_n}{n})\Big) \\ &= \frac{x_n}{n}(1 - \frac{x_n}{n})^{n-1} \\ &= \frac{x_n}{n} e^{-x_n} \end{aligned}$$

so that $n\Delta_n = nf(x_n) = x_n e^{-x_n}$. Suppose that we show that

$$2 - n^{-1} \leq x_n < 2 \tag{5}$$

with strict inequality on the left side for $n > 1$. Then, since $g(x) = xe^{-x} > 2e^{-2}$ for $1 \leq x < 2$, it follows that $n\Delta_n = nf(x_n) = g(x_n) \equiv x_n e^{-x_n} > 2\,e^{-2}$. Expanding $g(x)$ about $x = 2$ yields, with $x_n \leq x_n^* \leq 2$,

$$\begin{aligned} n\Delta_n &= g(x_n) = g(2) - g'(x_n^*)(2 - x_n) \\ &= 2e^{-2} - (1 - x_n^*)e^{-x_n^*}(2 - x_n) \\ &\leq 2e^{-2} + (2 - x_n)e^{-2} \\ &\leq e^{-2}(2 + n^{-1})\,. \end{aligned}$$

It remains to prove (5). It can be easily verified numerically for $n \leq 6$; see the table below. For the rest of the proof we assume $n > 6$. Now for $0 < x < 1$,

$$-1 - x^{-1}\log(1 - x) \leq \frac{1}{2}x + \frac{1}{3}x^2(1 - x)^{-1}$$

and hence it follows from (4) that with $y_n \equiv x_n/n$ we have

$$(n - 1)^{-1} \leq \frac{1}{2}y_n + \frac{1}{3}y_n^2(1 - y_n)^{-1}$$

which implies that

$$y_n \geq \frac{3(n+1)}{2(n-1)} \left\{ 1 - \left(1 - \frac{8(n-1)}{3(n+1)^2} \right)^{1/2} \right\} .$$

Now $1-(1-u)^{1/2} \geq (1/2)u(1+u/4)$, $0 < u < 1$, so that

$$\begin{aligned} y_n &\geq \frac{3(n+1)}{2(n-1)} \left\{ \frac{4(n-1)}{3(n+1)^2} \right\} \left\{ 1 + \frac{2(n-1)}{3(n+1)^2} \right\} \\ &= 2(n+1)^{-1} \left\{ 1 + \frac{n-1}{2n^2} + \epsilon_n \right\} \\ &> 2(n+1)^{-1} \left\{ 1 + \frac{n-1}{2n^2} \right\} \\ &= n^{-1}(2-n^{-1}) \end{aligned}$$

since

$$\epsilon_n \equiv \left\{ \frac{2(n-1)}{3(n+1)^2} - \frac{(n-1)}{2n^2} \right\} > 0$$

for $n > 6$. Thus $ny_n = x_n \geq (2-n^{-1})$. To prove the right inequality, assume it false for some $n \geq 3$. Then $y_n \geq 2/n$ and since the function g defined by $g(x) = -1 - x^{-1}\log(1-x) = x/2 + x^2/3 + \cdots$ is strictly increasing in $0 < x < 1$, it follows that

$$g(y_n) \geq g(2/n) = \sum_{k=1}^{\infty} (k+1)^{-1}(2/n)^k > \sum_{k=1}^{\infty} n^{-k} .$$

On the other hand, from (4)

$$g(y_n) = (n-1)^{-1} = \sum_{k=1}^{\infty} n^{-k} ,$$

a contradiction. Hence $x_n < 2$. The following table gives $n\Delta_n$ and the bounds for various n:

n	x_n	$n\Delta_n$	$(2+n^{-1})e^{-2}$
1	1.000	$e^{-1} = .3678\ldots$	.406
2	1.594	.324	.338
3	1.748	.304	.316
4	1.818	.295	.305
5	1.857	.290	.298
6	1.882	.287	.293
10	1.931	.280	.284
20	1.966	.275	.277
50	1.987	.272	.273
100	1.993	.272	.272
∞	2.	$.270671 = 2e^{-2} =$	.270671

3. PfS, Exercise 13.4.3, page 297:
(a) If $\phi''(0)$ is finite, then $\sigma^2 < \infty$.
(b) If $\phi^{(2k)}(0)$ is finite, then $EX^{2k} < \infty$.

Solution: (a) First, write

$$\phi'(t) = \lim_{h\to 0} \frac{\phi(t+h)-\phi(t)}{h} = \lim_{h\to 0} \frac{\phi(t)-\phi(t-h)}{h}$$

and

$$\begin{aligned}\phi''(0) &= \lim_{h\to 0} \frac{\phi'(h)-\phi'(0)}{h} = \lim_{h\to 0} \frac{\phi'(0)-\phi'(-h)}{h} \\ &= \lim_{h\to 0} \frac{\phi(h)-2\phi(0)+\phi(-h)}{h^2} \\ &= \lim_{h\to 0} E\left(\frac{e^{ihX}-2+e^{-ihX}}{h^2}\right) \\ &= -2\lim_{h\to 0} E\left(\frac{1-\cos(hX)}{h^2}\right) .\end{aligned}$$

But

$$\lim_{h\to 0} \frac{1-\cos(hx)}{h^2} = \frac{1}{2}x^2$$

and $(1-\cos(hx))/h^2 \geq 0$ for all x and h. Hence by Fatou's lemma

$$E(X^2) = 2E\left(\lim_{h\to 0}\frac{1-\cos(hX)}{h^2}\right) \leq 2\liminf_{h\to 0} E\left(\frac{1-\cos(hX)}{h^2}\right) = -\phi''(0) < \infty .$$

Thus $E(X^2) < \infty$ and $Var(X) < \infty$.

(b) The general case follows by induction on k as follows: Suppose that $\phi^{(2k-2)}(0)$ finite implies that $E(X^{2k-2}) < \infty$. Assume that $\phi^{(2k)}(0)$ is finite. Then $\phi^{(2k-2)}(t)$ exists and is continuous in a neighborhood of $t = 0$. By the induction hypothesis, $E(X^{2k-2}) < \infty$. By exercise 4.2 (a),

$$\phi^{(2k-2)}(t) = (-1)^{k-1} E(X^{2k-2} e^{itX}) .$$

If we define $H(x) \equiv E(X^{2k-2} 1_{[X \le x]})/E(X^{2k-2})$, then $H(x)$ is a d.f. with characteristic function

$$\psi(t) = \int e^{itx} dH(x) = \frac{(-1)^{k-1} \phi^{(2k-2)}(t)}{E(X^{2k-2})} .$$

Hence our hypothesis $\phi^{(2k)}(0)$ finite is equivalent to $\psi''(0)$ finite. From (a) it follows that

$$\frac{E(X^{2k})}{E(X^{2k-2})} = E_H(Y^2) \le -\psi''(0) < \infty ;$$

i.e. $E(X^{2k}) < \infty$.

4. Find independent random variables X, Y, and Z so that Y and Z do not have the same distribution, but $X+Y$ and $X+Z$ do have the same distribution.

   Solution: Let X be a rv with the de la Vallee - Poussin density $f(x) = (1-\cos(x))/\pi x^2$, and corresponding characteristic function $\phi(t) = (1-|t|) 1_{[-1,1]}(t)$. Let Z be a random variable with the charateristic function $\psi(t)$ which is ϕ extended periodically with period 4. (Draw the picture!) This corresponds to Z being a discrete rv with

$$p_k = P(Z = k\pi/2) = \frac{1}{4} \int_{-2}^{2} \psi(t) \exp(-i\pi kt/2) \, dt \qquad k \in \mathbb{Z} .$$

   Then, if $Y =_d X$ so that $\phi_Y = \phi$, we have, by independence

$$\phi_{X+Y}(t) = \phi_X \phi_Y = \phi^2 = \phi\psi = \phi_X \phi_Z = \phi_{X+Z}$$

   so that $X + Y =_d X + Z$, but Y and Z have different characteristic functions and hence also different distributions. (This is from Feller, *An Introduction to Probability Theory and Its Applications, Vol. II*, pages 505-507.)