Statistics 521, Problem Set 5 Solutions

Wellner; 11/2/2016

1. PfS, Exercise 3.4.2, page 48: Show that $\rho = 1$ if and only if $X - \mu_X = a(Y - \mu_Y)$ for some $a > 0$; and $\rho = -1$ if and only if $X - \mu_X = a(Y - \mu_Y)$ for some $a < 0$. Thus ρ measures linear dependence, not dependence.

   Solution: The "if" direction is easy: if $X - \mu_X = a(Y - \mu_Y)$ with $a > 0$, then $Cov(X, Y) = E(X - \mu_X)(Y - \mu_y) = aE(Y - \mu_Y)^2 = aVar(Y)$ and $Var(X) = a^2 Var(Y)$, which yields $\rho = 1$. Similarly, if $X - \mu_X = a(Y - \mu_Y)$ with $a < 0$, then $\rho = -1$. Conversely, suppose that $\rho^2 = 1$. Then $|Cov(X, Y)|^2 = Var(X)Var(Y)$, and hence, by the if and only if condition for equality in the Cauchy-Schwarz inequality, $\frac{|X-\mu_X|}{\sigma_X} = \frac{|Y-\mu_Y|}{\sigma_Y}$ a.s., or equivalently

$$|X - \mu_X| = \frac{\sigma_X}{\sigma_Y}|Y - \mu_Y|\,. \tag{1}$$

   To separate out what is going on in the two cases $\rho = 1$ and $\rho = -1$, consider first $\rho = 1$. Then we have equality throughout the system of inequalities given by

$$\begin{aligned} Cov(X, Y) &= E(X - \mu_X)(Y - \mu_Y) \\ &\le |E[(X - \mu_X)(Y - \mu_Y)]| \\ &\le E[|(X - \mu_X)(Y - \mu_Y)|] \\ &\le \sqrt{Var(X)Var(Y)}\,. \end{aligned}$$

   Equality in the third inequality implies that (**??**) holds. Equality in the second inequality implies that either $(X - \mu_X)(Y - \mu_Y) \ge 0$ a.s. or $(X - \mu_X)(Y - \mu_Y) \le 0$ a.s. (to see this, write out $|EY| = E|Y|$ in terms of positive and negative parts and use problem #3, problem set 4). But equality in the first inequality implies that $E\{[(X - \mu_X)(Y - \mu_Y)]^+\} \ge E\{[(X - \mu_X)(Y - \mu_Y)]^-\}$, and when combined with the preceding, this implies that $(X - \mu_X)(Y - \mu_Y) \ge 0$ a.s. Hence we conclude in this case that $(X - \mu_X)$ and $(Y - \mu_Y)$ have the same sign, and this in combination with (**??**) yields

$$X - \mu_X = \frac{\sigma_X}{\sigma_Y}(Y - \mu_Y)\,,$$

with $a = \frac{\sigma_X}{\sigma_Y} > 0$.

In the case $\rho = -1$, equality in the third inequality implies that (**??**) holds. But now we must have $-Cov(X, Y) = |Cov(X, Y)|$ which $E\{[(X - \mu_X)(Y - \mu_Y)]^-\} \geq E\{[(X - \mu_X)(Y - \mu_Y)]^+\}$, and when combined with the consequence of the second inequality (either $(X - \mu_X)(Y - \mu_Y) \geq 0$ a.s. or $(X - \mu_X)(Y - \mu_Y) \leq 0$ a.s.), this implies that $(X - \mu_X)(Y - \mu_Y) \leq 0$ a.s.; i.e. $(X - \mu_X)$ and $(Y - \mu_Y)$ have opposite signs. This in combination with (**??**) yields

$$X - \mu_X = -\frac{\sigma_X}{\sigma_Y}(Y - \mu_Y) = a(Y - \mu_Y)$$

with $a = -\frac{\sigma_X}{\sigma_Y} < 0$.

2. PfS, Exercise 3.4.3, page 48: (Littlewood's inequalities) Let $\mu_r \equiv E|X|^r$. For $r \geq s \geq t \geq 0$ we have $\mu_r^{s-t}\mu_t^{r-s} \geq \mu_s^{r-t}$. In particular, $\mu_2^3 \leq \mu_1^2\mu_4$.

   Solution: Note that we can rewrite the inequality as

   $$m_s \leq m_r^{(s-t)/(r-t)} m_t^{(r-s)/(r-t)}$$

   where $\alpha \equiv (r-t)/(s-t)$ and $\beta \equiv (r-t)/(r-s)$ satisfy $\alpha^{-1} + \beta^{-1} = (s-t)/(r-t) + (r-s)/(r-t) = 1$. Note that $r/\alpha + t/\beta = s$. Thus we see that Hölder's inequality with the powers α and β yields

   $$\begin{aligned} m_s = E|X|^s = E\{|X|^{r/\alpha}|X|^{t/\beta}\} &\leq \{E|X|^r\}^{1/\alpha}\{E|X|^t\}^{1/\beta} \\ &= m_r^{(s-t)/(r-t)} m_t^{(r-s)/(r-t)}. \end{aligned}$$

3. Suppose that $\epsilon_1, \ldots, \epsilon_n$ are i.i.d. random variables with $P(\epsilon_i = \pm 1) = 1/2$, and let $a_i \in \mathbb{R}$, $i = 1, \ldots, n$. Khintchine's inequalities say that for each $p > 0$

   $$A_p \left(\sum_{i=1}^n a_i^2\right)^{1/2} \leq \left(E|\sum_{i=1}^n a_i\epsilon_i|^p\right)^{1/p} \leq B_p \left(\sum_{i=1}^n a_i^2\right)^{1/2}.$$

   for some constants A_p and B_p. Prove the above inequalities when $p = 1$. **Hint:** The inequality on the right side is easy. Use the previous exercise to prove the inequality on the left side by showing that for $X \equiv \sum_{i=1}^n a_i \epsilon_i$, we have $E|X|^4 \le 3(E(X^2))^2$.

Solution: When $p = 1$, the upper bound is easy: by Liapunov's inequality (or by Jensen's inequality with $g(x) = x^2$), $E|X| \le (E|X|^2)^{1/2}$. Thus

$$E|\sum_1^n a_i \epsilon_i| \le \left(E|\sum_1^n a_i \epsilon_i|^2 \right)^{1/2} = \left(\sum_1^n a_i^2 \right)^{1/2},$$

so the upper bound holds with $B_1 = 1$.

For the lower bound, taking $r = 4$, $s = 2$, and $t = 1$ in the previous problem (Littlewood's inequalities) yields

$$E|X|^2 \le \{E|X|^4\}^{1/3} \{E|X|\}^{2/3},$$

or

$$\frac{\{E|X|^2\}^{3/2}}{\{E|X|^4\}^{1/2}} \le E|X|.$$

With $X = \sum_{i=1}^n a_i \epsilon_i$ we find that $E(X^2) = \sum_{i=1}^n a_i^2$ and

$$\begin{aligned} E|X|^4 &= E\left\{ \sum_{j,j',k,k'=1}^n a_j a_{j'} a_k a_{k'} \epsilon_j \epsilon_{j'} \epsilon_k \epsilon_{k'} \right\} \\ &= \sum_{i=1}^n a_i^4 + \binom{4}{2} \sum_{j<j'} a_j^2 a_{j'}^2 \\ &= \sum_{i=1}^n a_i^4 + 6 \sum_{j<j'} a_j^2 a_{j'}^2 \\ &\le 3 \left(\sum_{i=1}^n a_i^2 \right)^2 = 3\|a\|^4. \qquad (2) \end{aligned}$$

Hence it follows that

$$E|X| \ge \frac{\{E|X|^2\}^{3/2}}{\{E|X|^4\}^{1/2}} \ge \frac{(\sum a_i^2)^{3/2}}{\sqrt{3} \sum a_i^2} = \frac{1}{\sqrt{3}} \left(\sum_{i=1}^n a_i^2 \right)^{1/2}.$$

We conclude that Khintchine's inequality holds for $p = 1$ with $A_1 = 1/\sqrt{3}$ and $B_1 = 1$. The best possible constants A_p and B_p are known for all p; for $p = 1$ the best possible value of A_p is $1/\sqrt{2}$, and this is due to Szarek (1976), *Studia Math.* **63**, 197-208. For more on the case of general p and more general a_j's, see de la Peña and Giné (1999), *Decoupling*, pages 15-20 and 50.

Another way to get a bound of the same type as in (**??**) is via the sub-Gaussian exponential bound for X derived in class (where I called it Z):

$$P(|X| \geq t) \leq 2 \exp\left(-\frac{t^2}{2\|a\|^2}\right).$$

Putting this together with the formula $EY^r = \int_0^\infty r y^{r-1} P(Y \geq y) dy$ applied with $Y \equiv |Z|$ and $r = 4$ yields

$$\begin{aligned} E|X|^4 &= \int_0^\infty 4t^{4-1} P(|X| \geq t) dt \\ &\leq 8 \int_0^\infty t^3 \exp\left(-\frac{t^2}{2\|a\|^2}\right) dt \\ &= 4^2 \|a\|^4 \int_0^\infty y e^{-y} dy = 4^2 \|a\|^4 \end{aligned}$$

by the change of variables $y = t^2/(2\|a\|^2)$.

4. PfS, Exercise 3.5.3, page 55: Consider a probability measure P. (a) Let $Y \geq 0$ have df F. Show that $EY = \int_0^\infty P(Y \geq y) dy = \int_0^\infty [1 - F(y)] dy$. [Hint: prove the claimed formula for simple functions by summing by parts; and then the full claim follows from the MCT. A different proof to come later will use Fubini's theorem.]
(b) use the result of (a) to show that for $Y \geq 0$ and $\lambda \geq 0$ we have

$$\int_{[Y \geq \lambda]} Y \, dP = \lambda P(Y \geq \lambda) + \int_\lambda^\infty P(Y \geq y) dy.$$

Draw a picture to illustrate this.
(c) Suppose there is a $Y \in \mathcal{L}_1$ such that $P(|X_n| \geq y) \leq P(Y \geq y)$ for all $y > 0$ and all $n \geq 1$. Then use (b) to show that $\{X_n : \ n \geq 1\}$ is uniformly integrable.

Solution: (a) Suppose $Y \geq 0$. Then the simple functions

$$Y_n \equiv \sum_{k=1}^{n2^n} \frac{k-1}{2^n} 1_{[(k-1)2^{-n} \leq Y < k2^{-n}]}$$

satisfy $Y_n \nearrow Y$. For Y_n we compute

$$\begin{aligned}
EY_n &= \sum_{k=1}^{n2^n} \frac{k-1}{2^n} P((k-1)2^{-n} \le Y < k2^{-n}) \\
&= \sum_{k=1}^{n2^n} \left(\sum_{1}^{k-1} \frac{1}{2^n} \right) P((k-1)2^{-n} \le Y < k2^{-n}) \\
&= \sum_{k=1}^{n2^n} \sum_{1}^{n2^n} \frac{1}{2^n} 1_{[j \le k-1]} P((k-1)2^{-n} \le Y < k2^{-n}) \\
&= \sum_{j=1}^{n2^n} \frac{1}{2^n} \sum_{1}^{n2^n} 1_{[j \le k-1]} P((k-1)2^{-n} \le Y < k2^{-n}) \\
&= \sum_{j=1}^{n2^n} \frac{1}{2^n} P(j2^{-n} \le Y < n) = \sum_{j=1}^{n2^n} \frac{1}{2^n} P(j2^{-n} \le Y_n) \\
&= \int_0^\infty P(Y_n \ge y) dy
\end{aligned}$$

since $P(Y_n \ge n) = 0$ and $P(Y_n \in (j-1, j)/2^n) = 0$. By the MCT, the left side of the last display satisfies $EY_n \nearrow E(Y)$. Since $Y_n \nearrow$, we also have $P(Y_n \ge y) \nearrow P(Y \ge y)$ for each fixed y. Hence the right side of the last display satisfies

$$\int_0^\infty P(Y_n \ge y) dy \nearrow \int_0^\infty P(Y \ge y) dy$$

by the MCT again. Thus we conclude that

$$E(Y) = \int_0^\infty P(Y \ge y) dy = \int_0^\infty P(Y > y) dy = \int_0^\infty (1 - F(y)) dy$$

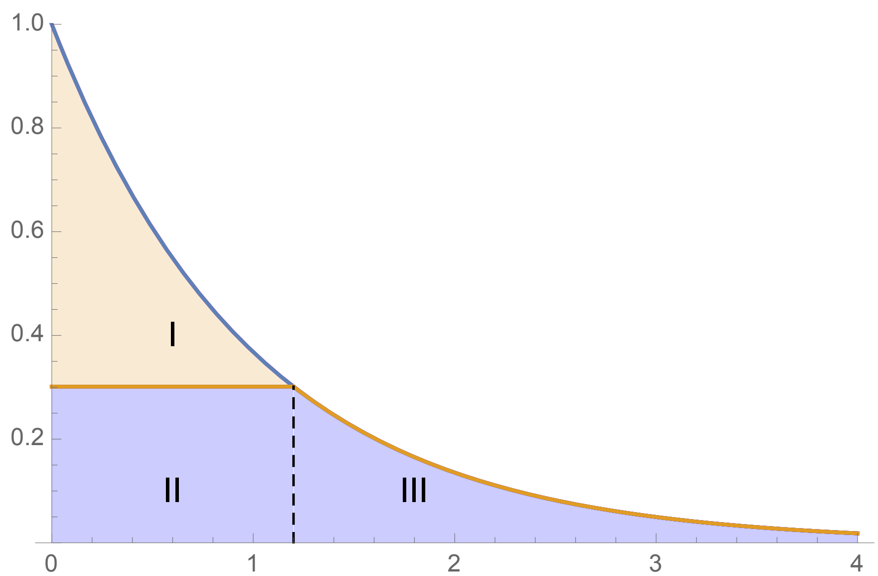
where the second equality holds since there are at most countably many points y with $P(Y > y) \ne P(Y \ge y)$.

(b) We will apply the formula in (a) to the random variable $Y = X1_{[X \ge \lambda]}$. Note that $Y = 0$ if $X < \lambda$, and $Y = X$ if $X \ge \lambda$. Hence we find that $P(Y = 0) = P(X < \lambda)$, and thus $P(Y > y) = 1 - P(X < \lambda)$

for $0 \leq y < \lambda$ while $P(Y > y) = 1 - P(X \leq y)$ for $\lambda \leq y < \infty$. Thus it follows from the formula in (a) that

$$\begin{aligned} E\{X1_{[X \geq \lambda]}\} &= E(Y) = \int_0^\infty P(Y > y)dy \\ &= \int_0^\lambda P(X \geq \lambda)dy + \int_\lambda^\infty P(X > y)dy \\ &= \lambda P(X \geq \lambda) + \int_\lambda^\infty P(X \geq y)dy\,; \end{aligned}$$

in the last equality we have used the fact that the number of discontinuities of $P(X \geq y)$ is at most countable, and hence of Lebesgue measure 0, and hence the two integrals are equal since the integrands differ on a set of Lebesgue measure at most 0. The following figure illustrates the identity: $EX1_{[X \geq \lambda]} = II_\lambda + III_\lambda$.



Contributions to $E(X)$: $III_\lambda = \int_\lambda^\infty P(X > y)dy$;
$II_\lambda = \lambda P(X > \lambda)$; $I_\lambda = \int_0^\lambda P(X > y)dy) - \lambda P(X > \lambda)$

(c) From (b) and the hypothesis it follows that

$$\begin{aligned} E\{|X_n|1_{[|X_n|\geq\lambda]}\} &= \lambda P(|X_n|\geq\lambda)+\int_\lambda^\infty P(|X_n|>y)dy \\ &\leq \lambda P(Y\geq\lambda)+\int_\lambda^\infty P(Y>y)dy \\ &= E\{Y1_{[Y\geq\lambda]}\} \end{aligned}$$

by (b) again in the last step. Thus we have

$$\sup_n E\{|X_n|1_{[|X_n|\geq\lambda]}\}\leq E\{Y1_{[Y\geq\lambda]}\}\to 0 \qquad \text{as } \lambda\to\infty;$$

i.e. $\{X_n\}$ is uniformly integrable.

5. (a) Show that if $|X_n|\leq Y$ and Y is integrable, then $\{X_n\}$ is uniformly integrable.
(b) Let $U\sim$ Uniform$(0,1)$, and let $X_n\equiv(n/\log n)1_{[0,1/n]}(U)$ for $n\geq 3$. Show that $\{X_n\}$ is uniformly integrable and $\int X_n dP\to 0$ even though they are not dominated by any integrable rv Y.
(c) Let $Z_n=n1_{[0,1/n]}(U)-n1_{[1/n,2/n]}(U)$. Show that $\{Z_n\}$ is not uniformly integrable, but that $\int Z_n dP\to 0$.

Solution: (a) Since $|X_n|\leq Y$ implies that $P(|X_n|\geq y)\leq P(Y\geq y)$, this follows from the preceding exercise.
(b) Now $X_n\geq 0$, $X_n\to_p 0\equiv X$, and $E(X_n)=1/\log(n)\to 0=E(X)$ as $n\to\infty$. Thus $\{X_n\}$ is uniformly integrable by Vitali's theorem. However the smallest rv above X_n for all $n\geq 3$ is the rv $Y=\sum_{k=3}^\infty \frac{k}{\log k}1_{(1/(k+1),1/k]}(U)$, and this has expectation

$$\begin{aligned} E(Y) &= \sum_{k=3}^\infty \frac{k}{\log(k)}\left\{\frac{1}{k}-\frac{1}{k+1}\right\} \\ &= \sum_{k=3}^\infty \frac{k}{\log(k)}\frac{1}{k(k+1)} \\ &= \sum_{k=3}^\infty \frac{1}{(k+1)\log(k)}=\infty\,. \end{aligned}$$

(c) Note that $E(Z_n) = 1 - 1 = 0 \to 0$, and $Z_n \to_p 0 \equiv Z$ since, for $\epsilon \leq 1$ we have $P(|Z_n| \geq \epsilon) = P(U \leq 2/n) = 2/n \to 0$. But

$$E|Z_n| = 2 \not\to 0 = E(Z)\,.$$

Hence by Vitali's theorem we conclude that $\{Z_n\}$ is not uniformly integrable.