

Dr. J. Hermon

This assignment is due in Canvas by 23:59 p.m. on Monday, September 29.**Please read the submission instructions on Canvas.*****Late assignments are not accepted.***

1. An ecology graduate student goes to a pond containing n water beetles and captures 60 of them, marks each with a dot of paint, and then releases them. A few days later she goes back and captures another sample of 50. Let X denote the number of marked beetles found in her sample of 50.
 - (a) Determine the probability mass function of X , i.e., find $P(X = k)$ for each k .
 - (b) She finds 12 beetles in her sample. Show that the function $P(X = 12)$ is initially an increasing function of n which then becomes decreasing after reaching a maximum value. Find the maximum likelihood estimate for n ; that is the value of n which maximises $P(X = 12)$.

Solution.

- (a) We observe the following.

- There are n choose 50 ways to choose the 50 waterbeetles.
- If we know that k of the 50 have marks, then those k can be chosen in 60 choose k ways, and the remaining $50 - k$ can be chosen in $n - 60$ choose $50 - k$ ways.
- The previous point helps us realize, that if $n - 60 < 50 - k$, there are not enough water beetles without marks for us to choose 50 with $50 - k$ not having marks.

Putting these together, we have:

$$P(X = k) = \begin{cases} \frac{\binom{60}{k} \binom{n-60}{50-k}}{\binom{n}{50}} & \text{if } n \geq 110 - k \\ 0 & \text{else} \end{cases}.$$

- (b) Let P_n be the probability that 12 marked beetles are in her sample of 50, if there were n beetles to begin with. The last bullet point in our observations in part (a) shows that $P_n(X = 12) = 0$ if $n \leq 98$. We start at $n = 98$:

$$P_{98} = \frac{\binom{60}{12} \binom{38}{38}}{\binom{98}{50}}.$$

Now, for $n > 98$, we compute:

$$\begin{aligned} P_n &= \frac{\binom{60}{12} \binom{n-60}{38}}{\binom{n}{50}} = \binom{60}{12} \frac{(n-60)!(n-50)!50!}{38!(n-98)!n!} \\ &= \frac{(n-60)(n-50)}{n(n-98)} \left(\binom{60}{12} \frac{((n-1)-60)!((n-1)-50)!50!}{38!((n-1)-98)!(n-1)!} \right) \\ &= \frac{(n-60)(n-50)}{n(n-98)} P_{n-1}. \end{aligned}$$

We can see, that if $\frac{(n-60)(n-50)}{n(n-98)} > 1$, then $P_n > P_{n-1}$. We test when this is the case. Note that, the following inequalities requires that $n(n-98) > 0$, which is the case when $n > 98$, but this is the case we are interested in.

$$\begin{aligned} \frac{n^2 - 110n + 60 \cdot 50}{n^2 - 98n} &> 1 \\ n^2 - 110n + 60 \cdot 50 &> n^2 - 98n \\ 60 \cdot 50 &> 12n \\ 250 &> n. \end{aligned}$$

So, we have that $P_{249} \geq P_{248}$, but $P_n < P_{n-1}$ for all $n \geq 250$. Therefore, the maximum likelihood estimate for n is 249.

2. By definition, a random variable is a function $X : \Omega \rightarrow \mathbb{R}$ such that $\{\omega : X(\omega) \leq x\} \in \mathcal{F}$ for every $x \in \mathbb{R}$. The distribution function $F_X : \mathbb{R} \rightarrow [0, 1]$ of the random variable X is defined by $F_X(x) = P(X \leq x)$. Suppose that X_1 and X_2 are both random variables.

- (a) Prove that $Y = \min\{X_1, X_2\}$ and $Z = \max\{X_1, X_2\}$ are also both random variables.
(b) Let X_1, X_2 be independent random variables. Find simple expressions for the distribution functions F_Y and F_Z in terms of F_{X_1} and F_{X_2} .

Solution.

- (a) We observe:

$$\begin{aligned}\omega \in \{Y \leq x\} &\Leftrightarrow \min\{X_1, X_2\}(\omega) \leq x \\ &\Leftrightarrow X_1(\omega) \leq x \vee X_2(\omega) \leq x \Leftrightarrow \omega \in \{X_1 \leq x\} \cup \{X_2 \leq x\}.\end{aligned}$$

Therefore, $\{Y \leq x\} = \{X_1 \leq x\} \cup \{X_2 \leq x\}$. Since $\{X_1 \leq x\}, \{X_2 \leq x\} \in \mathcal{F}$, and \mathcal{F} is a σ -algebra, we have that their union $\{X_1 \leq x\} \cup \{X_2 \leq x\} \in \mathcal{F}$. Since x was arbitrary, we conclude Y is a random variable.

We perform a similar calculation:

$$\begin{aligned}\omega \in \{Z \leq x\} &\Leftrightarrow \max\{X_1, X_2\}(\omega) \leq x \\ &\Leftrightarrow X_1(\omega) \leq x \wedge X_2(\omega) \leq x \Leftrightarrow \omega \in \{X_1 \leq x\} \cap \{X_2 \leq x\},\end{aligned}$$

which shows $\{Z \leq x\} = \{X_1 \leq x\} \cap \{X_2 \leq x\}$. The fact that σ -algebras are closed under intersections immediately affords $\{Z \leq x\} \in \mathcal{F}$.

- (b) Independence of X_1 and X_2 implies $P(\{X_1 \leq x\} \cap \{X_2 \leq x\}) = P(X_1 \leq x)P(X_2 \leq x)$. For Y , the inclusion exclusion principle gives us:

$$\begin{aligned}F_Y(x) &= P(Y \leq x) = P(\{X_1 \leq x\} \cup \{X_2 \leq x\}) \\ &= P(X_1 \leq x) + P(X_2 \leq x) - P(\{X_1 \leq x\} \cap \{X_2 \leq x\}) \\ &= F_{X_1}(x) + F_{X_2}(x) - P(X_1 \leq x)P(X_2 \leq x) \\ &= F_{X_1}(x) + F_{X_2}(x) - F_{X_1}(x)F_{X_2}(x).\end{aligned}$$

For Z , we simply apply independence:

$$F_Z(z) = P(Z \leq x) = P(\{X_1 \leq x\} \cap \{X_2 \leq x\}) = P(X_1 \leq x)P(X_2 \leq x) = F_{X_1}(x)F_{X_2}(x).$$

3. (a) Let Z_n be any sequence of random variables. Prove that if $P(|Z_n| \geq \varepsilon \text{ i.o.}) = 0$ for every $\varepsilon > 0$ then $P(\lim_{n \rightarrow \infty} Z_n = 0) = 1$.
(b) Let X_n be any sequence of random variables. Use part (a) and the Borel–Cantelli Lemma to prove that there are constants $c_n \rightarrow \infty$ such that $P(\lim_{n \rightarrow \infty} X_n/c_n = 0) = 1$.
Hint: Start by fixing a sequence $\varepsilon_n \downarrow 0$ and find c_n such that $P(|X_n| \geq \varepsilon_n c_n)$ is summable.

Solution.

- (a) Suppose that $P(|Z_n| \geq \varepsilon \text{ i.o.}) = 0$ for every $\varepsilon > 0$. Recall that $\{|Z_n| \geq \varepsilon \text{ i.o.}\}^c = \{|Z_n| < \varepsilon \text{ a.a.}\}$. Thus, $P(|Z_n| < \varepsilon \text{ a.a.}) = 1$ for every $\varepsilon > 0$.

Let $A_m = \{|Z_n| < \frac{1}{m} \text{ a.a.}\}$. We show that $A_m \searrow \{\lim_{n \rightarrow \infty} Z_n = 0\}$.

- If $\omega \in A_{m+1}$, then $|Z_n| \leq \frac{1}{m+1}$ for finitely many n , which implies that $|Z_n| \leq \frac{1}{m}$ for finitely many n , so $\omega \in A_m$. Therefore, $A_{m+1} \subset A_m$.
- Let $\omega \in \cap_m A_m$. Let $\varepsilon > 0$. Then, there exists some m such that $\frac{1}{m} < \varepsilon$. But then, $\omega \in A_m$, so that $|Z_n| < \frac{1}{m}$ for all but finitely many n . That is, there exists $N \in \mathbb{N}$ such that $|Z_n(\omega)| < \frac{1}{m} < \varepsilon$ for all $n > N$. Since this applies for all $\varepsilon > 0$, this implies $\lim_{n \rightarrow \infty} Z_n(\omega) = 0$. In conclusion, $\cap_m A_m \subseteq \{\omega : \lim_{n \rightarrow \infty} Z_n(\omega) = 0\}$.

Suppose $\omega \in \{\lim_{n \rightarrow \infty} Z_n = 0\}$. Let $m \in \mathbb{N}$. Then, there exists N such that for all $n > N$, we have $|Z_n(\omega)| < \frac{1}{m}$. In particular, $|Z_n(\omega)| < \frac{1}{m}$ almost always. Thus, $\omega \in A_m$. Since m was arbitrary, $\omega \in \cap_m A_m$. In conclusion, $\{\omega : \lim_{n \rightarrow \infty} Z_n(\omega) = 0\} \subseteq \cap_m A_m$. Therefore, $\{\omega : \lim_{n \rightarrow \infty} Z_n(\omega) = 0\} = \cap_m A_m$.

We apply continuity from above, to conclude that:

$$P(\lim_{n \rightarrow \infty} Z_n = 0) = P(\cap_m A_m) = \lim_{m \rightarrow \infty} P(A_m) = 1.$$

- (b) Let $\varepsilon_n = \frac{1}{n}$. Fix n . Now, $X_n \in [-N, N] \nearrow X_n \in \mathbb{R}$. By continuity of probabilities, we have $\lim_{n \rightarrow \infty} P(X_n \in [-N, N]) = P(X_n \in \mathbb{R}) = 1$. Therefore, there exists some N_n such that $P(X_n \in [-N_n, N_n]) \geq 1 - \frac{1}{2^n}$. In particular, this implies $P(|X_n| \geq N_n) \leq \frac{1}{2^n}$. Then, choose c_n such that $c_n \varepsilon_n \geq N$. In particular, $c_n > N \cdot n \geq n$. Therefore, $c_n \rightarrow \infty$.

In particular, this implies that

$$\begin{aligned} \sum_{n=1}^{\infty} P(|X_n| \geq c_n \varepsilon_n) &\leq \sum_{n=1}^{\infty} P(|X_n| \geq N_n) \text{ by monotonicity} \\ &\leq \sum_{n=1}^{\infty} \frac{1}{2^n} = 1. \end{aligned}$$

Now, fix $\varepsilon > 0$. Then, there exists some N such that $\varepsilon > \frac{1}{n}$ for all $n > N$. However, this implies that

$$\begin{aligned} \sum_{i=n}^{\infty} P\left(\frac{|X_i|}{c_i} \geq \varepsilon\right) &\leq \sum_{i=n}^{\infty} P\left(\frac{|X_i|}{c_i} \geq \frac{1}{i}\right) \text{ by monotonicity} \\ &= \sum_{i=n}^{\infty} P(|X_i| \geq c_i \varepsilon_i) \leq \sum_{n=1}^{\infty} P(|X_n| \geq c_n \varepsilon_n) \leq 1. \end{aligned}$$

By the Borel-Cantelli lemma, we have that $P\left(\frac{|X_n|}{c_n} \geq \varepsilon \text{ i.o.}\right) = 0$.

Since ε was arbitrary, this applies for every $\varepsilon > 0$, and we apply part (a) to conclude that $P\left(\lim_{n \rightarrow \infty} \frac{X_n}{c_n} = 0\right) = 1$. By construction, the constants $c_n \rightarrow \infty$.

4. Let $A_1, A_2, \dots \in \mathcal{F}$ and let τ be the tail field. If you have good intuition about the events $\{A_n \text{ i.o.}\}$ and $\{A_n \text{ a.a.}\}$ and about the meaning of τ then it should be “obvious” that these two events each lie in the tail field. Write a proof that $\{A_n \text{ i.o.}\} \in \tau$ and $\{A_n \text{ a.a.}\} \in \tau$.

Solution

First, we note that only need to prove $\{A_n \text{ a.a.}\} \in \tau$. Since σ algebras are closed under complements, $\sigma(A_{i_1}^c, A_{i_2}^c, \dots) = \sigma(A_{i_1}, A_{i_2}, \dots)$. So, a proof that $\{A_n \text{ a.a.}\} \in \tau$ will work just as well for $\{A_n^c \text{ a.a.}\} \in \tau$, and τ being a σ -algebra closed under complements gives us $\{A_n^c \text{ a.a.}\}^c = \{A_n \text{ i.o.}\} \in \tau$.

Now, we prove $A = \{A_n \text{ a.a.}\} \in \tau$. By definition, $\tau = \bigcap_{n \in \mathbb{N}} \sigma(A_n, A_{n+1}, \dots)$. Fix $n \in \mathbb{N}$. We show $A \in \sigma(A_n, A_{n+1}, \dots)$. First, we note $A = \bigcup_{m=1}^{\infty} \bigcap_{k=m}^{\infty} A_k$. However, since $\bigcap_{k=m}^{\infty} A_k \subseteq \bigcap_{k=m+1}^{\infty} A_k$ for every m , we have that $\bigcup_{m=1}^{\infty} \bigcap_{k=m}^{\infty} A_k = \bigcup_{m=n}^{\infty} \bigcap_{k=m}^{\infty} A_k$. However, this set is the intersection and union of sets exclusively in $\{A_n, A_{n+1}, \dots\}$, so we conclude that $A \in \sigma(A_n, A_{n+1}, \dots)$. Since n was arbitrary, this holds for all n , so $A \in \tau = \bigcap_{n=1}^{\infty} \sigma(A_n, A_{n+1}, \dots)$.

5. For integers $d \geq 2$, the integer lattice \mathbb{Z}^d is the set of all d -component vectors of integers. A bond is a pair $\{x, y\}$ of adjacent elements of \mathbb{Z}^d , i.e., with $\|x - y\|_1 = \sum_{i=1}^d |x_i - y_i| = 1$. In bond percolation, bonds are independently occupied with probability p and vacant with probability $1 - p$, where $p \in [0, 1]$ is fixed.

We enumerate the bonds in some manner, from inside out, as b_1, b_2, \dots . The sample space Ω consists of binary strings representing the occupancy status of the bonds, and the probability

space is the corresponding biased coin flip space (heads with probability p and tails with probability $1-p$ for fixed $p \in [0, 1]$), which is the modification of the probability space for independent fair coin flips (Section 2.6) with $P(A_{a_1 a_2 \dots a_n}) = p^h (1-p)^{n-h}$, where $h = \sum_{i=1}^n a_i$ denotes the number of heads among $a_1 a_2 \dots a_n$.

- (a) Let E be the event that there is an infinite connected cluster of occupied bonds. Prove that E truly is an event, i.e., is in the σ -algebra.
(Suggestion: for $x \in \mathbb{Z}^d$ and $k \in \mathbb{N}$, consider the event $E_{x,k}$ that x is a vertex in a connected cluster of size at least k vertices. Construct E from these events.)
- (b) Prove that E is in fact a tail event for the sequence of events O_1, O_2, \dots , where O_n is the event that the n th bond is occupied.
- (c) Let E_0 be the event that the origin is in an infinite cluster. Prove or disprove: E_0 is a tail event for the sequence of events O_1, O_2, \dots .

Solution.

- (a) We follow the hint. E occurs when an infinite cluster of occupied bonds exists. That is, there exists some vertex x such that x is in an infinite cluster. x being in an infinite cluster means that $\bigcap_{n=1}^{\infty} E_{x,n}$, since for every n the cluster it is a part of has at least n vertices because it is infinite. Since E is a statement saying that such an x must exist, it follows

$$E = \bigcup_{x \in \mathbb{Z}^d} \bigcap_{n=1}^{\infty} E_{x,n}.$$

- (b) Let Ω denote the set of grids. We can write $\Omega = \{0, 1\}^{\mathbb{N}}$, where the i -th element in the sequence is 1 if the i -th bond is occupied, and 0 if not. So, for $\omega \in \Omega$, we have $\omega = (\mathbf{1}_{O_1}(\omega), \mathbf{1}_{O_2}(\omega), \dots)$. Let π_i be the co-ordinate map. Now, consider the set Ω_n the set of grids on \mathbb{Z}^d with the first n bonds removed. Then define $\pi : \Omega \rightarrow \Omega_n$ as $\pi(\omega) = (\pi_n(\omega), \pi_{n+1}(\omega), \dots)$. In particular, by definition we have $O_i = \pi^{-1}(0, \dots, 0, 1, 0, \dots)$, where the 1 is the i -th slot. But, since the set of percolation events are on Ω_n are generated by $(1, 0, \dots), \dots, (0, \dots, 0, 1, 0, \dots)$, and unions and complements are preserved by the mapping π , we have that for any event $A \subset \Omega_n$, $\pi^{-1}(A) \in \sigma(O_n, \dots)$.

Also,

$$\pi^{-1}(A) = \{0, 1\} \times \dots \times \{0, 1\} \times A.$$

We show that $E = \pi^{-1}(\pi(E))$. Let $\omega \in E$. Fix $1 \leq i < n$. Suppose $\pi_i(\omega) = 1$. Define ω' such that $\pi_j(\omega) = \pi_j(\omega')$ for $j \neq i$, and $\pi_i(\omega') = 0$. If bond i is in an infinite cluster, and if i is contained in the path between any two elements in the infinite cluster (so that removing it breaks the cluster apart), then we must have that at least one of the two remaining clusters is infinite. Otherwise, the two clusters would be finite, and connecting them would result in another finite cluster, which contradicts the assumption that the original cluster was infinite. Therefore, $\omega' \in E$. If $\pi_i(\omega) = 0$, then defining ω' similarly to have $\pi_i(\omega') = 1$, would result in adding a bond, so clearly ω' 's infinite cluster would be undisturbed. Since this applies for all $1 \leq i < n$, we have that

$$E = \{0, 1\} \times \dots \times \{0, 1\} \times \pi(E) = \pi^{-1}(\pi(E)) \in \sigma(O_n, \dots).$$

Since this applies for all n , $E \in \tau$.

- (c) Fix Ω_2 from last problem, and consider that for the corresponding π , as we showed in the previous problem, $E \in \sigma(O_2, \dots)$ iff $E = \pi^{-1}(\pi(E))$, namely if $E = \{0, 1\} \times \pi(E)$. However, for every $\omega \in E$, we have $\pi(\omega) = 1$, with no $\omega' \in E$ such that $\pi(\omega') = 0$. Therefore, $E \neq \{0, 1\} \times \pi(E)$, so $E \notin \sigma(O_2, O_3, \dots)$, so it is not a tail event.

Reference book for percolation: G. Grimmett, *Percolation*, available for download at UBC library.

Recommended problems. The following problems from Rosenthal are recommended but are not to be handed in:

3.6.4 (assume that f is bounded below), 3.6.10, 3.6.12, 3.6.14 (assume the X_i are independent), 3.6.18 (This problem is different in different printings of the text; use $S_x = \{\liminf_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n \mathbf{1}_{A_i} \leq x\}$. If the A_i all have $P(A_i) = p \in (0, 1)$ then what do you think the graph of the function $f(x) = P(S_x)$ looks like? Later we'll have the theory for this, the Strong Law of Large Numbers.)

For solutions to even-numbered problems see: <http://www.probability.ca/jeff/grprobbook.html>.

Quote of the week:

I graduated from Douglass College without distinction. I was in the top 98% of my class and damn glad to be there. I slept in the library and daydreamed my way through history lecture. I failed math twice, never fully grasping probability theory. I mean, first off, who cares if you pick a black ball or a white ball out of the bag? And second, if you're bent over about the color, don't leave it to chance. Look in the damn bag and pick the color you want.

— Stephanie Plum in *Hard Eight* by Janet Evanovich