MATH 507a Homework 6

Qilin Ye

October 27, 2022

Problem: D3.1.2

If the X_i have a Poisson distribution with mean 1 then S_n has a Poisson distribution with mean n. Use Striling's formula to show that if $(k-n)/\sqrt{n} \to \mu$ then

$$\sqrt{2\pi n}\mathbb{P}(S_n = k) \to \exp(-x^2/2).$$

Proof. By Stirling,

$$\sqrt{2\pi n}\mathbb{P}(S_n = k) \sim \sqrt{2\pi n}\frac{e^{-n}n^k}{\sqrt{2\pi k}e^{-k}k^k} = e^{k-n}(n/k)^{1/2+k}.$$

Taking log, this becomes $k-n+(1/2+k)\log(n/k)$, where n/k can be re-written as $1-(k-n)/k \to 1-(x\sqrt{n})/k$. Then, using Taylor expansion $\log(1-t) = -t-t^2/2+o(t^2)$,

$$\begin{aligned} k-n+(1/2+k)\log(n/k) &\sim x\sqrt{n} + (1/2+k)\log(1-(x\sqrt{n}/k)) \\ &\sim x\sqrt{n} + (1/2+k)[-x\sqrt{n}/k + x^2n/(2k^2)] \\ &= -\frac{x\sqrt{n}}{2k} - \frac{x^2n}{2k} - \frac{x^2n}{4k^2}. \end{aligned}$$

Taking limits, only the second term survive with $x^2n/(2k) \to -x^2/2$, so $\sqrt{2\pi n}\mathbb{P}(S_n = k) \to \exp(-x^2/2)$. □

Problem: D3.2.2

Let $X_1, X_2, \ldots$ be independent with distribution F. Let $M_n = \max_{m\leqslant n} X_m$. Then $\mathbb{P}(M_n \leqslant x) = F(x)^n$. Prove the following limit laws for M_n:

(1) If $F(x) = 1-x^{-\alpha}$ for $x \geqslant 1$ where $\alpha > 0$, then for $y > 0$

$$\mathbb{P}(M_n/n^{1/\alpha} \leqslant y) \to \exp(-y^{-\alpha}).$$

(2) If $F(x) = 1-|x|^\beta$ for $-1 \leqslant x \leqslant 0$, where $\beta > 0$, then for $y < 0$,

$$\mathbb{P}(n^{1/\beta}M_n \leqslant y) \to \exp(-|y|^\beta).$$

(3) If $F(x) = 1-e^{-x}$ for $x \geqslant 0$, then for all y,

$$\mathbb{P}(M_n - \log n \leqslant y) \to \exp(-e^{-y}).$$

Proof. (1)
$$\mathbb{P}(M_n/n^{1/\alpha} \leqslant y) = \mathbb{P}(M_n \leqslant yn^{1/\alpha}) = (1 - 1/(ny^\alpha))^n \to \exp(-y^{-\alpha}).$$

(2) Similar to above,
$$\mathbb{P}(M_n n^{1/\beta} \leqslant y) = \mathbb{P}(M_n \leqslant yn^{-1/\beta}) = (1 - |y|^\beta/n)^n \to \exp(-|y|^\beta).$$

(3) Also similar to above,
$$\mathbb{P}(M_n - \log n \leqslant y) = \mathbb{P}(M_n \leqslant y + \log n) = (1 - \exp(-y - \log n))^n = (1 - e^{-y}/n)^n \to \exp(e^{-y}).$$
□

Problem: D3.2.3

Let $X_1, X_2, \ldots$ be i.i.d. standard normal distributions. We know
$$\mathbb{P}(X_i > x) \sim \frac{1}{\sqrt{2\pi}x} e^{-x^2/2}$$
as $x \to \infty$.

(1) Use this to conclude that for any real number θ,
$$\frac{\mathbb{P}(X_i > x + \theta/x)}{\mathbb{P}(X_i > x)} \to e^{-\theta}.$$

(2) Show that if we define b_n by $\mathbb{P}(X_i > b_n) = 1/n$ then
$$\mathbb{P}(b_n(M_n - b_n) \leqslant x) \to \exp(-e^{-x}).$$

Proof. (1) Using this identity,
$$\begin{aligned}\frac{\mathbb{P}(X_i > x + \theta/x)}{\mathbb{P}(X_i > x)} &= \frac{1/(\sqrt{2\pi}(x + \theta/x))}{1/(\sqrt{2\pi}x)} \frac{\exp(-(x + \theta/x)^2/2)}{\exp(-x^2/2)} \\ &= \frac{x + \theta/x}{x} \cdot \frac{\exp -(x^2 + 2\theta + \theta^2/x^2)/2}{\exp(-x^2/2)}.\end{aligned}$$

The first fraction $\to 1$ as $x \to \infty$, and in the second, $\theta^2/x^2 \to 0$, x^2 cancel each other out, and we are left with $\exp^{-2\theta/2}$. Therefore the ratio converges to $e^{-\theta}$.

Using (1), we have
$$\begin{aligned}\mathbb{P}(b_n(M_n - b_n) \leqslant x) &= \mathbb{P}(M_n \leqslant b_n + x/b_n) = [1 - \mathbb{P}(X_i > b_n + x/b_n)]^n \\ &= \left[1 - \mathbb{P}(X_i > b_n) \cdot \frac{\mathbb{P}(X_i > b_n + x/b_n)}{\mathbb{P}(X_i > b_n)}\right]^n \\ &\to (1 - e^{-x}/n)^n \to \exp(-e^{-x}).\end{aligned}$$
□

Problem: D3.2.12

Show that if $X_n \to c$ weakly where c is a constant, then $X_n \to c$ in probability.

Proof. By definition, $F_n \to F$ on all but possibly one point (namely c), since $F = 1_{[c,\infty)}$. Let $\epsilon > 0$ be given. Then

$$\mathbb{P}(|X_n - c| > \epsilon) = \mathbb{P}(X_n > c + \epsilon) + \mathbb{P}(X_n < c - \epsilon) = 1 - \mathbb{P}(X_n \leqslant c + \epsilon) + \mathbb{P}(X_n < c - \epsilon).$$

The second term $\to 1$ as $F_n(c+\epsilon) \to F(c+\epsilon) = 1$ and the third $\to 0$ similarly. Hence $\mathbb{P}(|X_n - c| > \epsilon) \to 0$, and $X_n \to c$ in probability, as claimed. ☐

Problem 1

For distribution functions F, G, define

$$\rho(F, G) := \inf\{\epsilon \geqslant 0 : F(x - \epsilon) - \epsilon \leqslant G(x) \leqslant F(x + \epsilon) + \epsilon \text{ for all } x \in \mathbb{R}\}.$$

(1) Show that ρ satisfies triangle inequality.

(2) Show that if $\rho(F_n, F) \to 0$ iff $F_n \to F$ weakly.

Proof. (1) Using the hint, let $\epsilon, \delta > 0$ be given; it suffices to show that if $\rho(F, G) < \epsilon, \rho(G, H) < \delta$, then $\rho(F, G) < \epsilon + \delta$.

Indeed, if $\rho(F, G) < \epsilon$ and $\rho(G, H) < \delta$, then

$$H(x) \leqslant G(x + \delta) + \delta \leqslant F(x + \epsilon + \delta) + \epsilon + \delta,$$

and similarly for the left half. That is, $\rho(F, H) \leqslant \epsilon + \delta$. Taking infimum over all eligible ϵ, δ, we obtain triangle inequality.

(2) We first assume $\rho(F_n, F) \to 0$ and let $\epsilon > 0$. There exists N such that if $n \geqslant N$ then $\rho(F_n, F) < \epsilon$, i.e., for all x,

$$F(x - \epsilon) - \epsilon \leqslant F_n(x) \leqslant F(x + \epsilon) + \epsilon.$$

Taking limits,

$$F(x - \epsilon) - \epsilon \leqslant \liminf_{n\to\infty} F_n(x) \leqslant \limsup_{n\to\infty} F_n(x) \leqslant F(x + \epsilon) + \epsilon.$$

Since ϵ is arbitrary, $F(x) \leqslant \liminf_{n\to\infty} F_n(x) \leqslant \limsup_{n\to\infty} F_n(x) \leqslant F(x)$, i.e., $F_n(x) \to F(x)$.

We now assume the converse, that $F_n \to F$ at all continuity points of F, and let $\epsilon > 0$ be given. Since discontinuity points are at most countable, we can pick continuity points $x_1 < \cdots < x_k$ such that $F(x_1) < \epsilon, F(x_k) > 1 - \epsilon$, and $x_i - x_{i-1} < \epsilon$ for terms in the middle. Since $F_n \to F$ on these points, and since there are finitely many terms that we pick, there exists N such that if $n \geqslant N$,

$$F(x_i) - \epsilon \leqslant F_n(x_i) \leqslant F(x_i) + \epsilon \text{ for all } i.$$

Now let x be given.

- If $x < x_1$, then $F_n(x) \leqslant F_n(x_1)$, and $F_n(x_1) \in [F(x_1) - \epsilon, F(x_1) + \epsilon]$, where $F(x_1) - \epsilon < 0$ and $F(x_1) + \epsilon < 2\epsilon$. Chaining everything together,

$$F_n(x) \leqslant F_n(x_1) \leqslant F(x_1) + \epsilon < 2\epsilon \leqslant F(x + 2\epsilon) + \epsilon,$$

and trivially

$$F(x-2\epsilon)-2\epsilon \leqslant F(x)-2\epsilon < 0 \leqslant F(x).$$

- The case $x > x_k$ is similar.
- Finally, if $x_{i-1} \leqslant x \leqslant x_i$ for some $2 \leqslant i \leqslant k$, by construction $x-\epsilon \leqslant x_{i-1} \leqslant x \leqslant x_i \leqslant x+\epsilon$, with distance $< \epsilon$ apart between each adjacent pair. This means $F_n(x) \leqslant F_n(x_i)$ and $F_n(x_i) < F(x_i)+\epsilon$, so chaining everything together,

$$F_n(x) \leqslant F_n(x_i) < F(x_i)+\epsilon \leqslant F(x+\epsilon)+\epsilon,$$

and similarly for the other side,

$$F(x-\epsilon)-\epsilon \leqslant F(x_{i-1})-\epsilon \leqslant F_n(x_{i-1}) \leqslant F_n(x).$$

Since ϵ is arbitrary we see $\rho(F, F_n) \leqslant 2\epsilon$ on its tail. Since ϵ is arbitrary we see $\rho(F, F_n) \to 0$. □

Problem 2

Show that if $\{X_n\}$ is tight and each $\mathbb{E}X_n < \infty$ with $\mathbb{E}X_n \to \infty$, then $\operatorname{var}(X_n) \to \infty$.

Proof. We use the hint and suppose $\mathbb{E}X_n \to \infty$ but $\operatorname{var}(X_n) \nrightarrow \infty$. That is, there exists a sequence X_{n_k} and $C \in [0, \infty)$ such that $\operatorname{var}(X_{n_k}) \leqslant C$ for all k, while $\mathbb{E}X_{n_k} \to \infty$. By Chevbyshev, for all M, and sufficiently large k (in particular $\mathbb{E}X_{n_k} \geqslant M$),

$$\begin{aligned}\mathbb{P}(|X_{n_k}| \leqslant M) &\leqslant \mathbb{P}(|X_{n_k} \leqslant M|) + \mathbb{P}(|X_{n_k}| \geqslant \mathbb{E}X_{n_k} + (\mathbb{E}X_{n_k} - M)) \\ &= \mathbb{P}(|X_{n_k} - \mathbb{E}X_{n_k}| \geqslant (\mathbb{E}X_{n_k} - M)) \\ &\leqslant \frac{\operatorname{var}(X_{n_k})}{(\mathbb{E}X_{n_k} - M)^2} \leqslant \frac{C}{(\mathbb{E}X_{n_k} - M)^2}.\end{aligned}$$

As $k \to \infty$, the denominator $\to \infty$, so $\mathbb{P}(|X_{n_k}| \leqslant M) \to 0$, meaning $\mathbb{P}(|X_{n_k}| > M) \to 1$, which shows $\{X_n\}$ is not tight. □

Problem 3

(1) Suppose $\mathbb{E}g(X) = \mathbb{E}g(Y)$ for all bounded continuous g. Show that X, Y have the same distribution.

(2) Suppose X, Y take values in $[0, 1]$ and $\mathbb{E}X^n = \mathbb{E}Y^n$ for $n \geqslant 1$. Show that X, Y have the same distribution.

Proof. (1) We define X_n be i.i.d. from X for each n. Then $\mathbb{E}g(X_n) = \mathbb{E}g(X) = \mathbb{E}g(Y)$, so by the characterization of weak convergence, $X_n \to Y$ in distribution, i.e., $\mathbb{P}(X_n \leqslant x) \to \mathbb{P}(Y \leqslant x)$ for all continuity points x of the d.f. of Y. But then this means the d.f. of X and that of Y agree on all continuity points, i.e., X and Y have the same distribution.

(2) Let $g \in C([0, 1])$ be (bounded) continuous and let ϵ be given. By the Weierstraß approximation theorem there exists a Bernstein polynomial $p \in C([0, 1])$ such that $\|p - g\|_\infty < \epsilon$. If we express $p(x)$ as $\sum_{j=0}^{n} a_j x^j$, then

from assumption

$$\mathbb{E}p(X) = \sum_{j=0}^{\infty} a_i \mathbb{E}X^j = \sum_{j=0}^{\infty} a_j \mathbb{E}Y^j = \mathbb{E}p(Y),$$

and

$$\mathbb{E}|g(X) - p(X)| = \int_{[0,1]} |g(X) - p(X)| \, \mathrm{d}\mathbb{P} < \epsilon$$

and similarly $\mathbb{E}|p(Y) - g(Y)| < \epsilon$. Therefore

$$\mathbb{E}|g(X) - g(Y)| \leqslant \mathbb{E}|g(X) - p(X)| + \mathbb{E}|p(X) - p(Y)| + \mathbb{E}|p(Y) - g(Y)| < 2\epsilon.$$

Since ϵ is arbitrary we see $\mathbb{E}g(X) = \mathbb{E}g(Y)$ for all bounded continuous g. By (1), X and Y have the same distribution. □

Problem 4

Consider i.i.d. random variables $\{X_n\}$ each equally likely to any of the values $a_1, \ldots, a_m$ uniformly. Let T_m be the number of trials needed to get some value to occur twice. That is, $T_m = \min\{n : X_n = X_\ell \text{ for some } \ell < n\}$. Note that if there are k trials, then the number of possible outcomes with all values distinct is $m(m-1)\cdots(m-k+1)$, and the total number of outcomes is m^k, so the probability of all k values are distinct is

$$\frac{m(m-1)\ldots(m-k+1)}{m^k} = \prod_{j=2}^{k}\left(1 - \frac{j-1}{m}\right).$$

Use this to show $\mathbb{P}(m^{-1/2}T_m > t) \to \exp(-t^2/2)$ as $m \to \infty$ for $t > 0$.

Proof. Since

$$\mathbb{P}(T_m > k) = \mathbb{P}(\text{no repeats in first } k) = \prod_{j=2}^{k}\left(1 - \frac{j-1}{m}\right),$$

we have

$$\begin{aligned}\log \mathbb{P}(m^{-1/2}T_m > t) &= \log \mathbb{P}(T_m > tm^{1/2}) \\ &= \sum_{j=2}^{tm^{1/2}} \log(1 - (j-1)/m).\end{aligned}$$

WLOG assume $m \geqslant 2$ and so $0 \leqslant (j-1)/m \leqslant 1/2$. Then using the identity $-x - x^2 \leqslant \log(1-x) \leqslant -x$, we have

$$\sum_{j=2}^{tm^{1/2}} \frac{j-1}{m} \leqslant -\log \mathbb{P}(m^{-1/2}T_m > t) \leqslant \sum_{j=2}^{tm^{1/2}} \left[\frac{j-1}{m} + \frac{(j-1)^2}{m^2}\right].$$

The first term is (assuming $tm^{1/2}$ is integer, since $x \sim \lfloor x \rfloor$)

$$m^{-1}tm^{1/2}(tm^{1/2} - 1)/2 \sim t^2/2$$

whereas the second has an extra

$$m^{-2}\sum_{j=2}^{tm^{1/2}} (j-1)^2 = m^{-2}\sum_{j=1}^{tm^{1/2}-1} j^2 \sim \frac{1}{6m^2}(tm^{1/2})(tm^{1/2})(2tm^{1/2} + 1) \sim \frac{C}{m^{-1/2}} \to 0.$$

Therefore, taking $m \to \infty$, $-\log \mathbb{P}(m^{-1/2}T_m > t)$ is sandwiched to $t^2/2$, or equivalently

$$\mathbb{P}(m^{-1/2}T_m > t) \to \exp(-t^2/2).$$

□