Lecture 4

Wed Jan/1/2024

Last time

- Weak convergence
- Levy's convergence Theorem.

Today

- Poisson Distribution
- Law of rare events

Poisson Distribution

Imagine we have a Bernoulli dist. $S_n = \sum_{k=1}^{n} X_k$ where X_i are iid r.v. with

$$\mathbb{P}(X_1 = 1) = p_n = 1 - \mathbb{P}(X_1 = 0).$$

Question: What happens to S_n as $n \to \infty$?

We can think of two regimes

- Fix probability: $p_n = p$. Then the CLT in HW 1 gives

$$\frac{S_n - np}{\sqrt{np(1-p)}} \xrightarrow{w} N(0,1).$$

- Fix the mean: $np_n = \lambda$. ← This models rare events!

  In this case the asymptotic dist

is Poisson!

Theorem: Suppose $S_n \sim \text{Binomial}(n, p_n)$, and $np_n \to \lambda \in (0, \infty)$ as $n \to \infty$. Then

$$\mathbb{P}(S_n = k) \longrightarrow e^{-\lambda} \frac{\lambda^k}{k!}.$$

We will use the place holder Poisson (λ) to denote this distribution.

We will prove a more general version of this result today. But first let's cover some properties of the Poisson dist.

Proposition The Poisson (λ) satisfies the following.

1) It's Moment generating func. is $e^{\lambda(e^t - 1)}$
2) It's characteristic func. is $e^{\lambda(e^{it} - 1)}$
3) It's mean is λ.
4) It's variance is λ.

Consequence sum of Poisson (λ_1) and Poisson (λ_2) is Poisson($\lambda_1 + \lambda_2$).

Proof:

1) By def

$$M(t) = \mathbb{E}(e^{tX}) = \sum_{k=0}^{\infty} e^{tk} \frac{e^{-\lambda}\lambda^k}{k!}$$

$$= e^{-\lambda} \sum_{k=0}^{\infty} \frac{(\lambda e^t)^k}{k!}$$

$$= e^{-\lambda} e^{e^t \lambda} = e^{\lambda(e^t - 1)}$$

2) Follows from $\varphi(t) = M(it)$.

3) Recall that $\mathbb{E}(X^p) = M^{(p)}(0)$. ← pth der.

$$M'(0) = \lambda e^t e^{\lambda(e^t - 1)} \Big|_0 = \lambda.$$

8) Similar to the above

$$M''(0) - (M'(0)) = \lambda.$$

↑ Exercise

□

Demo.

Law of rare events

The following is a more general version of the Theorem regarding Binomials.

Theorem (XX) For each n, let $X_{n,m}$ $\forall m \in [n]$ ($\{1, \dots, n\}$) be ind. r.v.'s such that

$$X_{n,m} = \begin{cases} 1 & \text{with prob } p_{n,m} \\ 0 & \text{with prob } 1 - p_{n,m}. \end{cases}$$

(Generalizes $p_{n_1} = \dots = p_{nn}$.)

Moreover, suppose that

1) $\sum_{m=1}^{n} p_{n,m} \to \lambda \in (0, \infty)$,

2 $\max_{m \in [n]} p_{n,m} \to 0$ as $n \to \infty$.

Let $S_n = \sum_{k=1}^{n} X_{nk}$, then,

$$S_n \xrightarrow{w} Z \quad \text{with} \quad Z \sim \text{Poisson}(\lambda).$$

To prove this result we will introduce a new notion of distance. For two measures μ and ν supported on a countable set S, we let the total variation distance be

$$\|\mu - \nu\|_{TV} = \frac{1}{2} \sum_{z} |\mu(z) - \nu(z)| = \sup_{A \subset S} |\mu(A) - \nu(A)|$$

will omit for simplicity (TV) — Exercise (=)

We will require a couple of Lemmas.

Lemma: The total variation distance defines a metric on prob. measures on $\mathbb{Z}$. Moreover, given a seq. $(\mu_n) \subseteq \mathcal{P}(\mathbb{Z})$ and $\mu \in \mathcal{P}(\mathbb{Z})$,

$$\mu_n \xrightarrow{w} \mu \iff \|\mu_n - \mu\| \to 0$$

Left as exercise.

Lemma If $\mu \times \nu \in \mathcal{P}(\mathbb{Z} \times \mathbb{Z})$ given by $(\mu \times \nu)(x,y) = \mu(x)\nu(y)$. Then

$$\|\mu_1 \times \nu_1 - \mu_2 \times \nu_2\| \leq \|\mu_1 - \mu_2\| + \|\nu_1 - \nu_2\|$$

Proof: By def

$$2\|\mu_1 \times \nu_1 - \mu_2 \times \nu_2\| = \sum_{x,y} |\mu_1(x)\nu_1(y) - \mu_2(x)\nu_2(y)|$$

$$\leq \sum_{x,y} |\mu_1(x)\nu_1(y) - \mu_2(x)\nu_1(y)| + \sum |\mu_2(x)\nu_1(y) - \mu_2(x)\nu_2(y)|$$

$$= \sum_y \nu_1(y) \sum_x |\mu_1(x) - \mu_2(x)| + \sum_x \mu_2(x) \sum_y |\nu_1(y) - \nu_2(y)|$$

$$\leq 2(\|\mu_1 - \mu_2\| + \|\nu_1 - \nu_2\|.$$

$\square$

Lemma If $\mu * \nu$ denotes the convolution

$$\mu * \nu(x) = \sum_y \mu(x-y)\nu(y).$$

Then, $\|\mu_1 * \nu_1 - \mu_2 * \nu_2\| \leq \|\mu_1 \times \nu_1 - \mu_2 \times \nu_2\|.$

Proof:

$$2\|\mu_1 * \nu_1 - \mu_2 * \nu_2\| = \sum_x \left|\sum_y \mu_1(x-y)\nu_1(y) - \mu_2(x-y)\nu_2(y)\right|$$

$$\leq \sum_x \sum_y |\mu_1(x-y)\nu_1(y) - \mu_2(x-y)\nu_2(y)|$$

$$= 2\|\mu_1 \times \nu_1 - \mu_2 \times \nu_2\|.$$

The sums are shifted, but $\mathbb{Z} - y = \mathbb{Z}$ $\forall y \in \mathbb{Z}$ □

Lemma: Let μ be such that $\mu(1) = p$ and $\mu(0) = 1-p$. Then,

$$\|\mu - \underbrace{\text{Poisson}(p)}_{\nu}\| \leq p^2.$$

Proof:

$$2\|\mu - \nu\| = |\mu(0) - \nu(0)| + |\mu(1) - \nu(1)| + \sum_{k \geq 2} \nu(k)$$

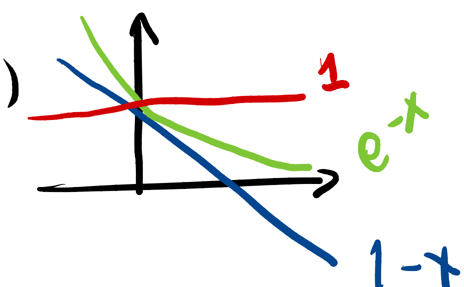
$$(\heartsuit) = |1 - p - e^{-p}| + |p - pe^{-p}| + 1 - e^{-p}(1+p).$$

Note that $1 - x \leq \exp(-x) \leq 1 \quad \forall x \geq 0$

$$(\heartsuit) = e^{-p} - 1 + p + p - pe^{-p} + 1 - e^{-p}(1+p)$$

$$= 2p(1 - e^{-p})$$

$$\leq 2p^2.$$



□

Proof of Theorem (x_x):

Let $\mu_{n,m}$ be the dist of $X_{n,m}$ an μ_n the dist of S_n. Let

$$\nu_{nm} = \text{Poisson}(p_{nm}), \quad \nu_n = \text{Poisson}\left(\sum_{m=1}^{n} p_{nm}\right)$$

and $\nu = \text{Poisson}(\lambda)$.

Since $\mu_n = \mu_{n1} * \dots * \mu_{nn}$ and $\nu_n = \nu_{n1} * \dots * \nu_{nn}$

the Lemmas above imply

$$\|\mu_n - \nu_n\| \le \sum_{m=1}^{n} \|\mu_{nm} - \nu_{nm}\| \le \sum_{m=1}^{n} p_{n,m}^2$$

This even give us a rate of convergence.

$$\le \max_{m \in [n]} p_{n,m} \sum_{m \in [n]} p_{n,m}$$

$$\to 0 \quad \text{as } n \to \infty.$$

Moreover, by construction $\nu_n \xrightarrow{w} \nu$.

Using triangle ineq.

$$\lim_{n\to\infty} \|\mu_n - \nu\| \le \lim_{n\to\infty} \|\mu_n - \nu_n\| + \|\nu_n - \nu\|$$

$$= 0.$$

Thus, by the first Lemma, $\mu_n \xrightarrow{w} \nu$. $\square$

We can easily generalize this result

Theorem 2.0 For each n, let $X_{n,m} \in \mathbb{N}$ $\forall m \in [n]$ be ind. r.v.'s such that

$$X_{n,m} = \begin{cases} 1 & \text{with prob } p_{n,m} \\ 0 & \text{with prob } 1 - p_{n,m} - \varepsilon_{n,m} \\ \ge 2 & \text{with prob } \varepsilon_{n,m} \end{cases}$$

Moreover, suppose that

1) $\sum_{m=1}^{n} p_{n,m} \to \lambda \in (0, \infty)$,

2) $\max_{m \in [n]} p_{n,m} \to 0$ as $n \to \infty$.

3) $\sum_{m=1}^{n} \varepsilon_{n,m} \to 0$ as $n \to \infty$.

Let $S_n = \sum_{k=1}^{n} X_{nk}$, then,

$$S_n \xrightarrow{w} Z \quad \text{with} \quad Z \sim \text{Poisson}(\lambda).$$

Proof: Define $X'_{nm} = 1$ if X_{nm}, and 0 otherwise. Theorem (**) gives that $S'_n = \sum_{m=1}^{n} X'_{nm}$ converges weakly to Z. Moreover

$$\|\mu_{S_n} - \mu_{S'_n}\| = \sum \|\mu_{X_{nm}} - \mu_{X'_{nm}}\|$$

$$\leq \sum (1 - p_{nm} - (1 - p_{nm} - \varepsilon_{nm})) + \varepsilon_{nm}$$

$$= 2 \sum \varepsilon_{nm}$$

$$\to 0.$$

□