

Section 4.4

Poisson approximation

- Recall, normal approximation: $\text{Bin}(n, p) \approx N(np, np(1 - p))$.
- Rule of thumb: Normal approx. is accurate when $np(1 - p) > 10$.

Q) What if “successes” are very rare?

A) Then $\text{Bin}(n, p)$ is approximately Poisson.

Def: A discrete r.v. X has Poisson distribution with parameter $\lambda > 0$ if

$$P(X = k) = \frac{e^{-\lambda} \lambda^k}{k!} \text{ for } k = 0, 1, 2, \dots$$

Notation: $X \sim \text{Poisson}(\lambda)$

Theorem: Let $\lambda > 0$, $k \in \{0, 1, 2, \dots\}$. Then

$$\lim_{n \rightarrow \infty} P\left(\text{Bin}\left(n, \frac{\lambda}{n}\right) = k\right) = P(\text{Poisson}(\lambda) = k) = \frac{e^{-\lambda} \lambda^k}{k!}$$

Precise meaning: Pmf of $\text{Bin}\left(n, \frac{\lambda}{n}\right)$ converges to pmf of $\text{Poisson}(\lambda)$.

Rough meaning: $\text{Bin}(n, p) \approx \text{Poisson}(np)$ for n large and p small.

Q) Why $\frac{\lambda}{n}$?

A) $\mathbb{E} \text{Bin}\left(n, \frac{\lambda}{n}\right) = \underline{n \cdot \frac{\lambda}{n}} = \lambda \checkmark \text{const.}$

Proof of special case: $k = 0$

$$P\left(\text{Bin}\left(n, \frac{\lambda}{n}\right) = 0\right) = \frac{\binom{n}{0} \cdot \left(\frac{\lambda}{n}\right)^0 \left(1 - \frac{\lambda}{n}\right)^n}{= \left(1 - \frac{\lambda}{n}\right)^n \xrightarrow{n \rightarrow \infty} e^{-\lambda}}$$

Fact:

$$\begin{aligned} \left(1 + \frac{1}{n}\right)^n &\xrightarrow{n \rightarrow \infty} e \\ \left(1 + \frac{\lambda}{n}\right)^n &\xrightarrow{n \rightarrow \infty} e^\lambda \\ \left(1 - \frac{\lambda}{n}\right)^n &\xrightarrow{n \rightarrow \infty} e^{-\lambda} \end{aligned}$$

$$P(\text{Poisson}(\lambda) = 0) = \frac{e^{-\lambda} \cdot \lambda^0}{0!} = e^{-\lambda} \quad \checkmark$$

Exercizi: Do other vals of k

Non-asymptotic approximation version

Theorem: Let $X \sim \text{Bin}(n, p)$, $Y \sim \text{Poisson}(np)$. Then for any subset $A \subset \{0, 1, 2, \dots\}$,

$$|P(X \in A) - P(Y \in A)| \leq np^2$$

Ex) Let $X \sim \text{Bin}\left(1000, \frac{1}{1000}\right)$. What is $P(X < 2)$?

$\approx \text{Poisson}(1)$

Exact solution: $P(X < 2) = \underbrace{P(0) + P(1)}_{a} = \binom{1000}{0} \cdot \left(\frac{1}{1000}\right)^0 \cdot \left(\frac{999}{1000}\right)^{1000} + \binom{1000}{1} \cdot \left(\frac{1}{1000}\right)^1 \cdot \left(\frac{999}{1000}\right)^{999}$

Poisson approx.: $P(X < 2) \approx \underbrace{\frac{e^{-1} \cdot 1^0}{0!} + \frac{e^{-1} \cdot 1^1}{1!}}_{b} = 2e^{-1} =: b$

$$\text{Thm: } |a - b| \leq np^2 = 1000 \cdot \frac{1}{1000^2} = \frac{1}{1000}$$

Properties of Poisson r.v.

Let $X \sim \text{Poisson}(\lambda)$. Then

1. $\mathbb{E}X = \lambda, \text{Var}(X) = \lambda$. } exercise to prove rigorously

Let $Y \sim \text{Poisson}(\alpha)$ be independent of X . Then

2. $X + Y \sim \text{Poisson}(\lambda + \alpha)$

Non-rigorous “proofs”:

1. $\mathbb{E}X = \mathbb{E}\text{Poisson}(\lambda) \approx \mathbb{E}\text{Bin}(\underline{n}, \underline{\frac{\lambda}{n}}) = n \cdot \frac{\lambda}{n} = \lambda$

$$\text{Var}(X) = \underline{\text{Var}(\text{Poisson}(\lambda))} \approx \text{Var}(\text{Bin}(n, \frac{\lambda}{n})) = n \frac{\lambda}{n} (1 - \frac{\lambda}{n}) = \lambda (1 - \frac{\lambda}{n}) \xrightarrow{n \rightarrow \infty} \lambda$$

2. $\text{Poisson}(\lambda) + \text{Poisson}(\alpha) \approx \text{Bin}(\lambda n, \frac{1}{n}) + \text{Bin}(\alpha n, \frac{1}{n})$

$$= \text{Bin}(\underline{\lambda n + \alpha n}, \underline{\frac{1}{n}}) \approx \text{Poisson}((\lambda + \alpha)n \cdot \frac{1}{n}) = \text{Poisson}(\lambda + \alpha) \quad \checkmark$$

Applications

Often it is natural to model as Poisson even without knowing n or p for underlying Binomial dist. (All we need is $\lambda = np = \mathbb{E}X$.)

Ex) An editor finds an average of 3 typos per page in a certain book. What is the probability that Page 74 has no typos? What about ≥ 3

typos? $\mathbb{E}X = 3$, where X is # of typos on arbitrary page
 $\lambda = 3$

$$X \sim \text{Poisson}(3)$$

$$P(\text{no typos}) = P(X=0) = e^{-3}$$

$$P(\geq 3 \text{ typos}) = P(X \geq 3) = 1 - P(X \in \{0, 1, 2\}) \\ = 1 - \left(e^{-3} + e^{-3} \cdot 3 + \frac{e^{-3} \cdot 3^2}{2} \right)$$

Ex) A factory has an average of 2 accidents per month. What is the probability of at least 3 accidents next month?

$X = \# \text{ of accidents next month} \sim \text{Poisson}(\lambda)$ $\lambda = 2$

$$\begin{aligned} P(\geq 3 \text{ accidents}) &= P(X \geq 3) \\ &= 1 - (P(0) + P(1) + P(2)) \\ &= 1 - \left[e^{-2} + e^{-2} \cdot 2 + \frac{e^{-2} \cdot 2^2}{2} \right] \end{aligned}$$