4.7. Poisson distribution

- Recall : $Y \sim Binom(n,p)$ if $Y = \#$ successes in n indep trials, where $p =$ prob. of success in each trial.

pmf: $p(k) = \binom{n}{k} p^k (1-p)^{n-k}$, $k = 0, 1, \dots, n$

$$E[Y] = np, \quad Var(Y) = np(1-p).$$

Ex In a class of 40 students, on average 2 students are sick. What is the prob. that 4 students are sick?

$X = \#$ sick students $\sim Binom(\underset{n}{\underset{\parallel}{40}}, p)$. $E[X] = 40 \cdot p = 2 \Rightarrow p = 1/20$.

$$p(4) = \binom{40}{4} \left(\tfrac{1}{20}\right)^4 \left(1 - \tfrac{1}{20}\right)^{36} = 0.09012$$

Approximation to Binomial

- The pmf $p(k)$ of Binom(n,p) is sometimes difficult to compute, especially for large n.

- $\Rightarrow$ approximate. Assume that successes are rare (like in the last example):

$$\underbrace{n \to \infty}_{\text{many trials}}, \quad \underbrace{np \to \lambda = const}_{\text{const \# of successes on ave.}} \qquad (\text{thus } p \to 0)$$

- Let us approximately compute $p(k)$: $(p \simeq \lambda/n)$

$$p(k) \simeq \frac{n!}{k!(n-k)!} \left(\frac{\lambda}{n}\right)^k \left(1 - \frac{\lambda}{n}\right)^{n-k}$$

$$= \frac{n(n-1)\cdots(n-k+1)}{k!} \frac{\lambda^k}{n^k} \left(1-\frac{\lambda}{n}\right)^n \left(1 - \frac{\lambda}{n}\right)^{-k}.$$

Note: (1) $\frac{n(n-1)\dots(n-k+1)}{n^k} \to 1$ $(n \to \infty, k \text{ fixed})$

(2) $\left(1 - \frac{\lambda}{n}\right)^n \to e^{-\lambda}$ $(n \to \infty)$ [Recall: $\lim_{x\to\infty} \left(1 + \frac{x}{n}\right)^n = e^x$ for each $x \in R$]

(3) $\left(1 - \frac{\lambda}{n}\right)^{-k} \to 1$ $(n \to \infty, k \text{ fixed})$.

Hence

$$\boxed{p(k) \simeq \frac{\lambda^k}{k!} e^{-\lambda}}$$

Def (Poisson distr.)

A r.v. X has the Poisson distribution with parameter λ $(\lambda>0)$ if X has pmf

$$p(k) = e^{-\lambda}\frac{\lambda^k}{k!}, \quad k=0,1,2,\dots$$

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

We proved:

Thm (Poisson approximation) "Poisson paradigm"

Let $Y \sim \text{Binomial}(n,p)$. If $n\to\infty$, $np\to\lambda=\text{const}$, then the pmf of $Y \longrightarrow$ pmf of Poisson(λ):

$$p_Y(k) \to e^{-\lambda}\frac{\lambda^k}{k!} \quad \text{for each } k=0,1,2,\dots$$

Cor If $X \sim \text{Poisson}(\lambda)$ then

$$E[X]=\lambda, \quad Var(X)=\lambda$$

Proof: By Thm, X is approximately $\sim \text{Binom}(n,p)$ where $n\to\infty$, $np\to\lambda$.
Recalling the expected value and the variance of binomial distr,

$$E[X]\approx np=\lambda, \quad Var(X)=np(1-p)\approx np \approx \lambda, \quad (\text{since } p\to 0)$$

QED

Examples of Poisson distr: # of misprints on a page of a book
# of software failures on a given day;
# of accidents on a section of a highway on a given day (here $n=\infty$!)

Ex In Ex. p.40, $X\approx\text{Poisson}(2)$. $\Rightarrow$ $p(4)\approx e^{-2}\cdot\frac{2^4}{4!}\approx \underline{0.09022}$ (compare to 0.09012)

Ex On average 2 accidents occur on campus on a given day.
(a) $P\{\text{no accidents}\}=?$ (b) $P\{\text{at least one accident}\}=?$ (c) $P\{\text{at least two}\}=?$

$X=\text{Poisson}(2)$ (since $\lambda=E(X)=2$)

(a) $p(0)=e^{-2}=0.14$, (b) $1-p(0)=1-e^{-2}=0.87$

(c) $1-p(0)-p(1)=1-e^{-2}-2e^{-2}=0.56$.

Remark When can one use Poisson approx?
Ans: When the number of trials is large (or infinite), # successes expected to be const
E.g. $n=100$, $k\le 5$ or: # of accidents on a highway (trials = every instant)

Remark: $1=\sum_{k=0}^{\infty}p(k)=\sum_{k=0}^{\infty}e^{-\lambda}\frac{\lambda^k}{k!} \Rightarrow \boxed{\sum_{k=0}^{\infty}\frac{\lambda^k}{k!}=e^{\lambda}}$ (We gave a probabilistic "proof" of this classical identity)

People v. Collins

From Wikipedia, the free encyclopedia

The People of the State of California v. Collins[1] was a 1968 jury trial in California, USA that made notorious forensic use of mathematics and probability.

Contents

Trial at first instance

Bystanders to a robbery in Los Angeles testified that the perpetrators had been a black male, with a beard and moustache, and a caucasian female with blonde hair tied in a ponytail. They had escaped in a yellow motor car.

After testimony from an "instructor in mathematics" about the multiplication rule for probability, the prosecutor invited the jury to consider the probability that the accused pair, who fitted the description of the witnesses, were not the robbers. Even though the "instructor" had not discussed conditional probability, the prosecutor suggested that the jury would be safe in estimating:

Black man with beard	1 in 10
Man with moustache	1 in 4
White woman with pony tail	1 in 10
White woman with blonde hair	1 in 3
Yellow motor car	1 in 10
Interracial couple in car	1 in 1000

Chance that a couple meets these characteristics is

$$\frac{1}{10}\cdot\frac{1}{4}\cdot\frac{1}{10}\cdot\frac{1}{3}\cdot\frac{1}{10}\cdot\frac{1}{1000}=\frac{1}{12,000,000}$$

The jury returned a verdict of guilty.

Appeal

The Supreme Court of California set aside the conviction, criticising the statistical reasoning and disallowing the way the decision was put to the jury. In their judgment, the justices observed that mathematics:

> *... while assisting the trier of fact in the search of truth, must not cast a spell over him.*

- Fallacies of prosecution: (a) the characteristics are not independent! (e.g. beard & moustache), Hence the probabilities can not be multiplied

  (b) Even if the characteristics were independent, 1:12,000,000 would be the probability to come across such a couple at random (without search).

- Correcting (b): compute Prob. that there is another couple in LA with these characteristics.

  5,000,000 couples in LA

  $\Rightarrow$ $\underbrace{\text{Average \# of couples with these char's}}_{= X}$ is $\frac{5}{12} =: \lambda$.

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

  $$P\{X \geq 2 \mid X \geq 1\} = \frac{P\{X \geq 2\}}{P\{X \geq 1\}} = \frac{1 - e^{-\lambda} - \lambda e^{-\lambda}}{1 - e^{-\lambda}} \approx 0.19$$

  Not so small.