MATH 180A (Lecture A00)

mathweb.ucsd.edu/~ynemish/teaching/180a

Today: Expectation

Next: ASV 3.4

Week 5:

- Homework 3 due Friday, February 10
- Regrades of Midterm 1, HW 1, HW2 active on Gradescope until February 12, 11 PM

Rare events. Poisson distribution

$\forall x \in \mathbb{R} \quad \sum_{k=0}^{\infty} \frac{x^k}{k!} = e^x$

Let $\lambda > 0$ and let X be a r.v. taking values in $\{0,1,2,\dots\}$.
X has Poisson distribution with parameter λ if

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

We write $X \sim \mathrm{Pois}(\lambda)$

Poisson distribution describes the probability that a "rare" event occurs k times after repeating the experiment (independent trials) "many" times.

Is this a probability distribution?

$$P(X=k) \geq 0, \quad \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} e^{-\lambda} = e^{-\lambda} \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} = e^{-\lambda} e^{\lambda} = 1$$

λ gives the "expected number" of occurrances

Rare events. Poisson distribution

$$\lim_{n\to\infty}\left(1+\frac{x}{n}\right)^n = e^x$$

Let X be the number of successes in n independent trials with success probability $\frac{\lambda}{n}$, $\lambda > 0$.

Then $P(X=k) = \binom{n}{k}\left(\frac{\lambda}{n}\right)^k\left(1-\frac{\lambda}{n}\right)^{n-k}$

What happens if $n \to \infty$ ($k \in \{0,1,2,\dots\}$ is fixed)?

$$\lim_{n\to\infty}\binom{n}{k}\left(\frac{\lambda}{n}\right)^k\left(1-\frac{\lambda}{n}\right)^{n-k} = \lim_{n\to\infty}\frac{n!}{k!\,(n-k)!}\,\frac{\lambda^k}{n^k}\left(1-\frac{\lambda}{n}\right)^n\left(1-\frac{\lambda}{n}\right)^{-k}$$

$$= \frac{\lambda^k}{k!}\cdot\lim_{n\to\infty}\frac{n\cdot(n-1)(n-2)\cdots(n-k+1)}{n^k}\left(1-\frac{\lambda}{n}\right)^n\left(1-\frac{\lambda}{n}\right)^{-k}$$

$n\cdot n\cdot n \cdots n$

$$= \frac{\lambda^k}{k!}\lim_{n\to\infty} 1\cdot\underbrace{\left(1-\cancel{\frac{1}{n}}\right)}_{1}\underbrace{\left(1-\cancel{\frac{2}{n}}\right)}_{1}\cdots\underbrace{\left(1-\cancel{\frac{(k-1)}{n}}\right)}_{1}\left(1-\frac{\lambda}{n}\right)^n\underbrace{\left(1-\cancel{\frac{\lambda}{n}}\right)^{-k}}_{1} = \frac{\lambda^k}{k!}e^{-\lambda}$$

Poisson distribution. Example

Observation: between 1875 and 1894 (20 years) in 14 units of Prussian army there were 196 deaths from horse kicks, distributed in the following way

# deaths per unit per year, k	# unit-years with k deaths	empirical probability	$P(X=k)$
0	144	0 51	
1	91	0.33	
2	32	0.11	
3	11	0.04	
4	2	0.01	
5+	0	0	

total 280

$P(X=0)=e^{-0.7}$

$P(X=1)=0.7\,e^{-0.7}$

$P(X=2)=\frac{(0.7)^2}{2}e^{-0.7}$

Let is "expected number" of death per unit

Poisson distribution. Example

A 100 year storm is a storm magnitude expected to occur in any given year with probability $\frac{1}{100}$. Over the course of a century, how likely is it to see at least 4 100 year storms?

$S_{n,p}$

We can compute this as $P(S_{100,\frac{1}{100}} \geq 4)$ for $S_{100,\frac{1}{100}} \sim Bin(100, \frac{1}{100})$

$$P(S_{100,\frac{1}{100}} = k) = \binom{100}{k}\left(\frac{1}{100}\right)^k\left(1-\frac{1}{100}\right)^{100-k} \approx \frac{\lambda^k}{k!}e^{-\lambda} = \frac{e^{-1}}{k!}$$

$$P(S_{100,\frac{1}{100}} \geq 4) = \sum_{k=4}^{100} \binom{100}{k}\left(\frac{1}{100}\right)^k\left(1-\frac{1}{100}\right)^{100-k} = 1 - \sum_{k=0}^{3}\binom{100}{k}\left(\frac{1}{100}\right)^k\left(1-\frac{1}{100}\right)^{100-k}$$

$$\approx 1 - \sum_{k=0}^{3}\frac{e^{-1}}{k!} = 1 - \frac{1}{e}\left(1 + 1 + \frac{1}{2} + \frac{1}{6}\right)$$

Summary

Independent trials: the most important (discrete) probability distributions are:

- $\mathrm{Ber}(p)$: $P(X=1)=p$, $P(X=0)=1-p$, $0 \le p \le 1$

  (single trial with success probability p)

- $\mathrm{Bin}(n,p)$: $P(S_n=k)=\binom{n}{k} p^k (1-p)^{n-k}$, $0 \le k \le n$

  (number of successes in n independent trials with rate p)

- $\mathrm{Geom}(p)$: $P(N=k)=(1-p)^{k-1} p$, $k=1,2,3,\ldots$

  (first successful trial in repeated independent trials with rate p)

- $\mathrm{Poisson}(\lambda)$: $P(X=k)=e^{-\lambda}\frac{\lambda^k}{k!}$, $k=0,1,2,\ldots$ $\lambda>0$

  (approximates $\mathrm{Bin}(n,\frac{\lambda}{n})$, number of rare events in many trials)

Expectation

Example Toss a fair coin 1000 times, and record the sequence of outcomes 1 1 0 0 1 0 0 1 1 0 1 0 0

Average then $\frac{1}{1000}(1+1+0+0+1+0+0+1+1 \cdots) \approx \frac{1}{2} = \frac{1}{2}\cdot 1 + \frac{1}{2}\cdot 0$

What if the coin is biased $P(X_j=1)=p$, $P(X_j=0)=1-p$?

Then the average (random) is approximately $p = p\cdot 1 + (1-p)\cdot 0$

Def. Let X be a discrete random variable with possible values $t_1, t_2, t_3 \ldots$. The expectation (or expected value, or mean) of X is

weigted average

$$E(X) := \sum_j t_j \cdot P(X = t_j)$$

Expectation

Q: Is the expectation $E(X)$ the value that X is equal to most often?

(a) Yes, always

(b) No, not generally

Example Let X be the number rolled on a fair die.

$$E(X) = \sum_{k=1}^{6} k \cdot P(X=k) = \sum_{k=1}^{6} k \cdot \frac{1}{6} = \frac{1}{6}(1+2+3+4+5+6) = \frac{21}{6} = \frac{7}{2}$$

Example Let Y be $Ber(p)$.

$$E(Y) = 0 \cdot (1-p) + 1 \cdot p = p$$

Expectation

Example You toss a biased coin repeatedly until the first heads. How long do you expect it to take?

N = the time the first heads comes up, $N = \text{Geom}(p)$

$$E(N) = \sum_{k=1}^{\infty} k \cdot P(N=k) = \sum_{k=1}^{\infty} k \cdot (1-p)^{k-1} \cdot p$$

$$\overset{k-1=\ell}{=} p \sum_{\ell=0}^{\infty} (1+\ell)(1-p)^{\ell} = p \cdot \sum_{\ell=0}^{\infty} (1-p)^{\ell} + p \sum_{\ell=0}^{\infty} \ell (1-p)^{\ell}$$

$$= p \cdot \frac{1}{1-(1-p)} + (1-p) \sum_{\ell=1}^{\infty} \ell (1-p)^{\ell-1} p = 1 + (1-p) E(N)$$

$$E(N) = 1 + (1-p) E(N), \quad E(N)(1-(1-p)) = 1, \quad E(N) = \frac{1}{p}$$