Stat 134 lec 17

warmup 11:00–11:10

Caroline has a busy email: new messages arrive with rate 9 per hour, and she checks her email with rate 1 per hour, independent of email arrivals. Suppose both follow Poisson arrival processes.

Suppose there is no new message now, what is the chance that she checks her email before any new message arrives?

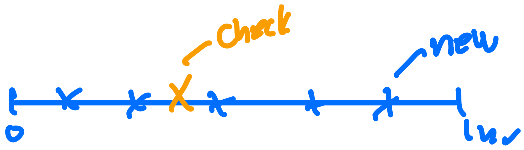
related question!

If roll a die 100 times what is the chance you get a 1 before you get a 2 or 4? — guess its $1/3$

Focus on times you get a 1, 2 or 4, call a 1 a success and 2 or 4 a failure. Think of a single Bernoulli trial with 1 a success and 2 or 4 a failure, the chance the Bernoulli trial is a success is $1/3$.

In original problem we have the superposition of 2 indep Pois processes $Pois(\mu=9) + Pois(\mu=1)$

$= Pois(\mu=10)$



Suppose you get 15 arrivals (either new or check). Just focus on these 15 bernoulli trials, the chance the 1st one is a check is $1/10$.

Announcements

For Wednesday review, write down questions in discussion board on b-course by Tuesday 8pm.

Last time

sec 3.5 Poisson distribution

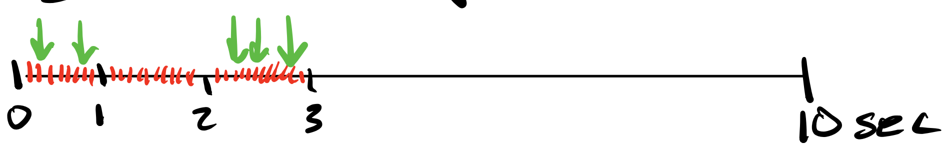
$X \sim Pois(\mu)$

$$P(X=k) = \frac{e^{-\mu}\mu^k}{k!}$$

$E(X) = Var(X) = \mu.$

Poisson Process or Poisson Random Scatter (PRS):

eg radioactive decay of Americium 241 in 10 seconds



Assumptions

1. no two particles arrive at the same time.
   (this allows us to divide 10 sec into n small time intervals each with at most one arrival.
2. X is a sum of n ← large iid Bernoulli(p) trials.
   ↖ small

$\mu = np$ is avg # of arrivals in 10 sec.
$\lambda = \mu/10$ is the arrival rate per second.

X = # arrivals in 10 seconds.
Suppose $\lambda = 4$ arrivals/sec
then $\mu = \lambda \cdot 10 = 40 \Rightarrow X \sim \text{Pois}(40)$
Americium has a long half life.
Y = # arrivals in 12070 sec.
$Y \sim \text{Pois}(\lambda \cdot 12070)$
4

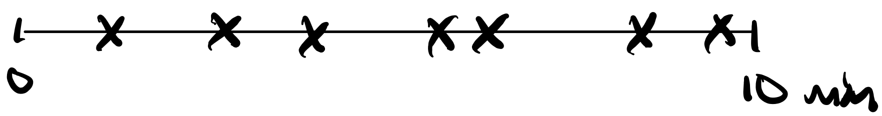
Today ① Finish section 3.5 Poisson Thinning

② midterm review

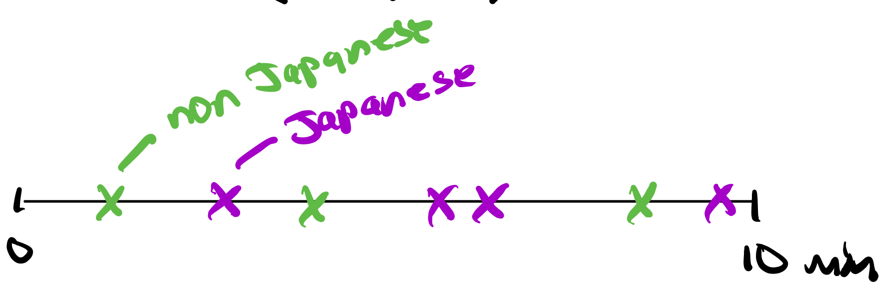
Sec 3.5 Poisson thinning

Cars arrive at a toll booth according to a Poisson process at a rate $\lambda = 3$ arrivals/min

X = # cars arriving at a toll booth in 10 min. $X \sim \text{Pois}(\underbrace{\lambda \cdot 10}_{30})$



Of cars arriving, it is known, over the long term, that 60% are Japanese imports.



Call Japanese cars a success and non Japanese a failure.

$$\# \text{ cars} \sim \text{Pois}(\lambda \cdot 10) = \text{Pois}(30)$$

$$\# \text{ Japanese imports} \sim \text{Pois}(p\lambda \cdot 10) = \text{Pois}(18)$$

$$\# \text{ non Japanese} \sim \text{Pois}(q\lambda \cdot 10) = \text{Pois}(12)$$

Ex What is the probability that in a given 10 min interval, 15 cars arrive at the booth and 10 are Japanese imports?

$X = \#$ cars in 10 min $\sim Pois(30)$

indep

$J = \#$ Japanese cars in 10 min $\sim Pois(18)$

$nJ = \#$ nonJapanese $\sim Pois(12)$

$P(X=15, J=10) = P(nJ=5, J=10)$

$= P(nJ=5)P(J=10)$

$= \boxed{\dfrac{e^{-12}\,12^5}{5!} \cdot \dfrac{e^{-18}\,18^{10}}{10!}}$ ✓

$= P(J=10 \mid X=15) \cdot P(X=15)$

↑ $Bin(15, 3/5)$

$$\binom{15}{10} \cdot \left(\frac{3}{5}\right)^{10} \left(\frac{2}{5}\right)^{5} \qquad e^{-30} \frac{30^{15}}{15!}$$

② midterm review

Which distributions are (approximately) a sum of a fixed number of independent Bernoulli trials?

Discrete

name and range	$P(k) = P(X=k)$ for $k \in$ range	mean	variance
uniform on $\{a, a+1, \ldots, b\}$ $\{1,2,\ldots,n\}$	$\frac{1}{b-a+1}$	$\frac{a+b}{2}$ $\frac{n+1}{2}$	$\frac{(b-a+1)^2-1}{12}$ $\frac{n^2-1}{12}$
✓ Bernoulli (p) on $\{0,1\}$	$P(1)=p;\ P(0)=1-p$	p	$p(1-p)$
✓ binomial (n,p) on $\{0,1,\ldots,n\}$	$\binom{n}{k}p^k(1-p)^{n-k}$	np	$np(1-p)$
✓ Poisson (μ) on $\{0,1,2,\ldots\}$	$\frac{e^{-\mu}\mu^k}{k!}$	μ	μ
✓ hypergeometric (n, N, G) (N G n) on $\{0,\ldots,n\}$ if $n \ll N$	$\frac{\binom{G}{k}\binom{N-G}{n-k}}{\binom{N}{n}}$	$\frac{nG}{N}$	$n\left(\frac{G}{N}\right)\left(\frac{N-G}{N}\right)\left(\frac{N-n}{N-1}\right)$
geometric (p) on $\{1,2,3\ldots\}$	$(1-p)^{k-1}p$	$\frac{1}{p}$	$\frac{1-p}{p^2}$
geometric (p) on $\{0,1,2\ldots\}$	$(1-p)^k p$	$\frac{1-p}{p}$	$\frac{1-p}{p^2}$
negative binomial (r,p) on $\{0,1,2,\ldots\}$	$\binom{k+r-1}{r-1}p^r(1-p)^k$	$\frac{r(1-p)}{p}$	$\frac{r(1-p)}{p^2}$
✓ normal	$\phi(x)$	μ	σ^2

approx binomial for p not too small or large and n large

or By CLT normal is sum of iid Bernoulli trials

DeMorgan's rule: $(A \cap B)^c = A^c \cup B^c$

$\Rightarrow A \cap B = (A^c \cup B^c)^c$

so $\boxed{P(A \cap B) = 1 - P(A^c \cup B^c)}$

Inclusion exclusion formula:

Let A_1, A_2, A_3 be dependent RVs with $P(A_i) = .9$ for $i = 1, 2, 3$.

Find a lower bound for $P\left(\bigcap_{i=1}^{3} A_i\right)$

$P\left(\bigcap_{i=1}^{3} P_i\right) = 1 - P\left(\bigcup_{i=1}^{3} A_i^c\right)$ by DeMorgan's rule.

$$P\left(\bigcup_{i=1}^{3} A_i^c\right) = P(A_1^c) + P(A_2^c) + P(A_3^c)$$

$$- P(A_1^c A_2^c) - P(A_2^c A_3^c) - P(A_1^c A_3^c)$$

$$+ P(A_1^c A_2^c A_3^c)$$

$$\Rightarrow P\left(\bigcup_{i=1}^{3} A_i^c\right) \le P(A_1^c) + P(A_2^c) + P(A_3^c) = 3(.1) = .3$$

$$P\left(\bigcap_{i=1}^{3} P_i\right) \ge 1 - .3 = .7$$

ex Conditional distribution, Poisson

8. Let X_1 and X_2 be independent random variables such that for $i = 1, 2$, the distribution of X_i is Poisson (μ_i). Let m be a fixed positive integer. Find the distribution of X_1 given that $X_1 + X_2 = m$. Recognize this distribution as one of the famous ones, and provide its name and parameters.

$$\left.\begin{array}{l} X_1 \sim \text{Pois}(\mu_1) \\ X_2 \sim \text{Pois}(\mu_2) \end{array}\right\} \text{ indep} \qquad P(X_1 = k) = \frac{e^{-\mu_1}\mu_1^k}{k!}$$

$X_1 \mid X_1 + X_2 = m$ takes value $0, 1, 2, \dots, m$

$$P(X_1 = k \mid X_1 + X_2 = m) = \frac{P(X_1 = k, X_2 = m-k)}{P(X_1 + X_2 = m)}$$

$$= \frac{P(X_1 = k)\,P(X_2 = m-k)}{P(X_1 + X_2 = m)}$$

$$= \frac{\dfrac{\cancel{e^{-\mu_1}}\mu_1^k}{k!} \cdot \dfrac{\cancel{e^{-\mu_2}}\mu_2^{m-k}}{(m-k)!}}{\dfrac{\cancel{e^{-(\mu_1+\mu_2)}}(\mu_1+\mu_2)^m}{m!}}$$

$$= \binom{m}{k}\left(\frac{\mu_1}{\mu_1+\mu_2}\right)^k \left(\frac{\mu_2}{\mu_1+\mu_2}\right)^{m-k}$$

$$\Rightarrow X_1 \mid X_1 + X_2 = m \sim \text{Bin}\left(m, \frac{\mu_1}{\mu_1+\mu_2}\right)$$

Problem 4 (conditional probability)

Two jars each contains r red marbles and b blue marbles. A marble is chosen at random from the first jar and placed in the second jar. A marble is then randomly chosen from the second jar. Find the probability this marble is red.

X = the first marble color

Y = the second marble color.

R = red

B = blue

$$\frac{r+1}{r+b+1} \cdot \frac{r}{r+b}$$

$$P(Y=R) = \underbrace{P(Y=R \mid X=R)P(X=R)}_{= \frac{r+1}{r+b+1} \cdot \frac{r}{r+b}} + \underbrace{P(Y=R \mid X=B)P(X=B)}_{= \frac{r}{r+b+1} \cdot \frac{b}{r+b}}$$

expectation question

An urn contains 90 marbles, of which there are 20 greens, 20 blacks and 50 red marbles. Tom draw marble without replacement until the 6^{th} green marble. Let X = # of marbles drawn. Example: **GGG**BRB**GG**BR**G** with $x = 11$. Find $\mathbb{E}[X]$.

Hint First find the expected number of marbles until the 1^{st} green marble.
What is the min and max of X? — 1, 71

$P = 1/21$

$$X = I_1 + \dots + I_{70} + 1$$

where $I_2 = \begin{cases} 1 & \text{if } 2^{nd} \text{ nongreen is before } 1^{st} \text{ green} \\ 0 & \text{else} \end{cases}$

2^{nd} nongreen must go here

$\_\, G_1 \,\_\, G_2 \,\_ \;\dots\; \_\, G_{20} \,\_$

$P = 1/21$ since there are 21 slots the 2^{nd} nongreen can go, $\Rightarrow E(X) = 70 \cdot (1/21) + 1$

To solve original problem
by symmetry you expect # nongreen marbles between first and second green also to be $70 \cdot (1/21) + 1$. Similar between 2^{nd} and 3^{rd} green etc

Hence answer is $\boxed{6 \cdot \left[70 (1/21) + 1\right]}$