Stat 134 lec 17

Warmup 1:00-1:10

A class of 60 students includes 20 seniors. For a group project, the class is split at random without replacement into 10 groups of 6 students each. Find the expected number of groups that contain no seniors.

$X = I_1 + \cdots + I_{10}$

$p = \frac{\binom{20}{0}\binom{40}{6}}{\binom{60}{6}}$

$I_i = \begin{cases} 1 & \text{if } i^{\text{th}} \text{ group has no seniors} \\ 0 & \text{else} \end{cases}$

$\boxed{E(X) = 10 \cdot p}$

Announcements

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

Midterm

Friday 1:08 — 2pm + 10 min to upload to gradescope.

Question 1: Copy Cal honor code.

5 or 6 probability questions

leave answers unsimplified.

no calculator

Closed book, closed notes.

will provide distribution summary table (p476 Pitman)

Today ① midterm review

① 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

By CLT — sum of a large number of iid RVs is approx normal,

$I_1 + \cdots + I_n \approx N(np, npq)$

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

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

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 $P\left(\bigcap_{i=1}^{3} A_i\right) \geq$

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

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

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

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

inclusion exclusion rule.

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

$\Rightarrow P\left(\bigcap_{i=1}^{3} A_i\right) \geq \boxed{1 - 3(.1)}$

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 $0, 1, 2, \ldots, m$

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

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

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

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

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

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

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

$$\parallel$$

$$P(Y=R) = P(Y=R \mid X=R)P(X=R)$$

$$+ P(Y=R \mid X=B)P(X=B)$$

$$\parallel$$

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

ex 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 $\mathbb{E}[X]$.

ex Chebyshev

You are using a telescope to measure the speed at which the planet Saturn crosses the night sky. To do this you draw two lines on your lens, and measure the time it takes for Saturn to cross between the two lines. However, your time measurement is noisy, so you will conduct this observation several times and average their results.

Let X_i represent the time measurement from the *ith* observation. Your measurements are well calibrated, so for each i, $E(X_i) = \mu_X$, where μ_X is the true time it takes Saturn to cross between the lines. Each measurement also has standard deviation $SD(X_i) = 0.03$ seconds.

a You will take n measurements, $X_1, \ldots, X_n$, using the same procedure, and use the sample average $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ to estimate μ_X. In terms of n, what is $SD(\bar{X})$?

b What is the smallest number of measurements you will need to take so that your estimate $\bar{X}$ has at most a $\frac{1}{25}$ probability of falling outside the interval $\mu_X \pm 0.003$ seconds? (Hint: Chebyshev) ex 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. $X \sim \text{Pois}(\underbrace{\lambda \cdot 10}_{30})$

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

$$P(X=k) = \frac{e^{-30}\, 30^{k}}{k!}$$