3Y03 - 3J04 PROBABILITY & STATISTICS FOR (CIVIL) ENGINEERING

C01 - Lecture 10

Last time

BINOMIAL DISTRIBUTION : X = # successes in n trials

$$f(x) = \binom{n}{x} p^x (1-p)^{n-x}$$

GEOMETRIC DISTRIBUTION: X = # trials till 1st success

$$f(x) = p(1-p)^{x-1}$$

special case with $r=1$.

NEGATIVE BINOMIAL DISTRIBUTION : X = # trials till rth success

$$f(x) = \binom{x-1}{r-1} p^r (1-p)^{x-r}$$

HYPERGEOMETRIC DISTRIBUTION : X = # successes, sampling from N items (containing K 'success' items) WITHOUT REPLACEMENT

Binomial: like sampling n times WITH replacement

→ prob. of picking a "success item" at each step stays the same.

Hypergeometric : sampling WITHOUT replacement means probability of picking a "success item" at each step changes depending on what came before.

$$\underset{\uparrow f(x)}{P(X=x)} = \frac{\binom{K}{x}\binom{N-K}{n-x}}{\binom{N}{n}}$$

for $\max\{0, n-(N-K)\} \le x \le \min\{n, K\}$

Usually $n \le K$ so this is just $0 \le x \le K$.

We could decompose X into a sum of other random variables but lack of replacement makes that more complicated & hence calculations messier.

we would get $\mu = np$ where p = prob. of choosing a success from the pool of items i.e. $p = \frac{K}{N}$

and $\sigma^2 = np(1-p)\underbrace{\left(\frac{N-n}{N-1}\right)}$ ← "finite population correction factor"

If n small relative to N then we can approximate hypergeometric here with Binomial $\left(n, \frac{K}{N}\right)$ (easier to compute!)

<u>Example</u> $\overset{N}{100}$ chips, $\overset{(N-K)}{75}$ conforming, $\overset{K}{25}$ non-conforming

$\overset{n}{3}$ sampled without replacement $X = \#$ non-conforming

What is P (at most 2 fail to conform)?

<u>Solution</u>
$$P(X \leq 2) = \sum_{x=0}^{2} \frac{\binom{K}{x}\binom{N-K}{n-x}}{\binom{N}{n}}$$

$$= \sum_{x=0}^{2} \frac{\binom{25}{x}\binom{75}{3-x}}{\binom{100}{3}} \quad \ldots\ldots = 0.99$$

(see Lecture 4)

Also $P(X \le 2) = 1 - P(X=3) = 1 - \dfrac{\binom{25}{3}\binom{75}{0}}{\binom{100}{3}}$

$= 1 - 0.01 = 0.99.$

<u>Example</u> Lotto 6/49: You pick 6 #s from $\{1, \ldots, 49\}$

6 drawn in lottery. X = # matches between <u>your 6</u> & drawn 6 ↓ n

Find $f(x) = P(X=x)$.

<u>Solution</u> Here "success" = # from the drawn 6 K

"failure" = # from the other 43 not drawn $N-K$

$N=49$

$$\frac{\binom{K}{x}\binom{N-K}{n-x}}{\binom{N}{n}}$$

x	$f(x)$
0	$\binom{6}{0}\binom{43}{6} / \binom{49}{6}$
1	$\binom{6}{1}\binom{43}{5} / \binom{49}{6}$
$\vdots$	

and so on. (up to $x=6$, because you can have at most 6 matches, of course.)

<u>Poisson Distribution</u>

"Incidents" take place randomly during a period of time T

or "Flaws" distributed randomly on a region of size T

where we know on average λ # incidents / unit time

or λ # flaws / unit length/area

X = # incidents or # flaws

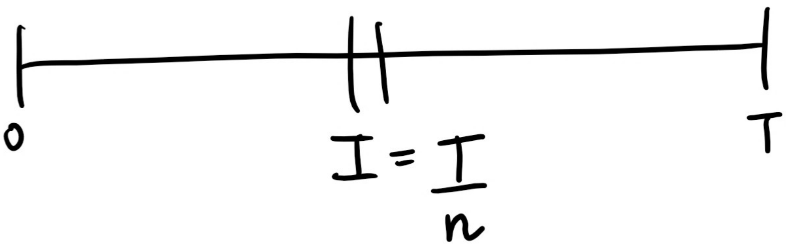
e.g. # earthquakes in 5 years

flaws along a pipeline of 10m

buses to show up in 1hr (if schedule is broken) e.g. in the U.K. :)

Notice if X has Poisson distribution with parameters T, λ, then $\mu = E(X) = \lambda T$.

(We'll be able to see this "formally" below once we workout $f(x)$:))



Divide T into n pieces. Choose n so big that prob. of >1 incident/flaw in region of size $\frac{T}{n}$ is negligible.

Prob. of incident/flaw in one of these pieces $\cancel{\approx}$ is $\frac{\lambda T}{n}$.

Call this p.

If incidents/flaws are independent, can model this with Binomial distribution, parameters n, p. (For each of the n little regions, there is a prob. p that that region has an incident/flaw.)

With this approximation (some choice of n):

$$f(x) = P(X=x) = \binom{n}{x} p^x (1-p)^{n-x}$$

$$= \frac{n!}{(n-x)!\,x!} \left(\frac{\lambda T}{n}\right)^x \boxed{\left(1 - \frac{\lambda T}{n}\right)^{n-x}} \to e^{-\lambda T}$$

$$\xrightarrow[n \to \infty]{} \boxed{\frac{(\lambda T)^x}{x!} e^{-\lambda T}}$$

Poisson : approximated by Binomial (better approx. for bigger n) so we let $n \to \infty$ to get the true distribution (the limit of all the approximations)

$$E(X) = \mu = \lambda T \quad , \quad V(X) = \lambda T.$$

$= \sum_{x=0}^{\infty} x \left(\frac{(\lambda T)^x}{x!}\right) e^{-\lambda T}$ → for calculations, see textbook.

We say that " X is a Poisson random variable with mean λT "

Which unfortunately is sometimes, including on the Formula Sheet, called just λ instead of μ, in which case $f(x) = \frac{\lambda^x}{x!} e^{-\lambda x}$.

<u>Example</u> Next time.