Last time (Session 9)
Chebyshev, CLT, ∞ sum rule, 
craps principle

This time (Session 10)
Geometric and Negative Binomial 
(collector's problem)
Poisson scatters

Recall (Problem 3.1.10)

\[ \text{Independent} \]

\[ \text{Bin} (n=72, p=\frac{1}{36}) + \text{Bin} (n=54, p=\frac{1}{36}) = \text{Bin} (n=126, p=\frac{1}{36}) \]

\[ \Rightarrow \text{Indep. } \]

\[ \text{Pois} (\mu_1=2) + \text{Pois} (\mu_2=1.5) = \text{Pois} (\mu=3.5) \]

So, sums of 
independent Poissons 
are Poisson.

Proof

\[ N_1 \sim \text{Poisson} (\mu_1=2) \]
\[ N_2 \sim \text{Poisson} (\mu_2=1.5) \]

\[ P(N_1+N_2=5) = \sum_{j=0}^{5} P(N_1=j) P(N_2=5-j) \]

\[ = \sum_{j=0}^{5} \left[ e^{-2} \frac{2^j}{j!} \right] \left[ e^{-1.5} \frac{(1.5)^{5-j}}{(5-j)!} \right] \]

\[ = e^{-3.5} \frac{(3.5)^5}{5!} \times 2 \]

\[ \sum_{j=0}^{5} \frac{5!}{j!(5-j)!} \left( \frac{2}{3.5} \right)^j \frac{(1.5)^{5-j}}{(3.5)^{5-j}} \]

\[ = 1 = \sum \text{Bin. Prob.'s} \]
Example: Lump of carbon
\[ N = \# \text{Carbon }^{14} \text{ atoms} = 5 \text{ zillion} \]
\[ p = \text{prob. of decay} = \frac{1}{\text{1 zillion}} \text{ atom; in next minute} \]
so
\[ N_1 = \# \text{decays in next minute} \sim \text{Poisson} (\mu = \text{1}) \]

Lump of Strontium
\[ N_2 = \# \text{decays in next minute} \sim \text{Poisson} (\mu = 3) \]
Then
\[ N_1 + N_2 \sim \text{Poisson} (\mu = 4) \]

Examples of Approx. Poisson r.v.'s
# decays per minute in lump containing several radioactive elts
# fatal car accidents in Tippecanoe Co.
in one week
# deaths by horse-kick per year
in a Prussian Cavalry Corp.
# misprints on book page

Recall earlier claim (Session 8)
\[ \sum_{k=1}^{n} N = X_1 + X_2 + \ldots + X_n \]
\[ \text{independent}, X_k \sim \text{Bernoulli} (p_k) \]
\[ \text{all } p_k \text{'s small} \]
then
\[ N \approx \text{Poisson} (\mu = \sum_{k=1}^{n} P_k) \]

Decay ex. cont.
Given that total # decays in one min. is 4,
What is the (cond.) prob. that exactly 3 were \(^{14}\text{C}\) decays (and one was Stron.)?
\[ = \frac{P(N_1 = 3 \mid N_1 + N_2 = 4)}{P(N_1 + N_2 = 4)} = \frac{e^{-\frac{3}{2}} \left( \frac{3}{2} \right)^3}{e^{-\frac{3}{2}} \left( \frac{3}{2} \right)^3} \]
\[ = \frac{1}{3!} \left( \frac{5}{8} \right)^3 \left( \frac{3}{8} \right)^2 \]

Given that
\[ N_1 + N_2 = n, \]
\[ N_1 \sim \text{Binomial} (n, p = \frac{\lambda_1}{\lambda_1 + \lambda_2}) \]
Poisson Scatter

Random pattern of "hits"
in space (or time)

Examples

1-D: Times when phone calls come into switchboard. (Section 4.2)

2-D: Points at which raindrops hit sidewalk over 1 second interval.

---

Poisson Scatter Theorem

If

1. No multiple hits (at 1 point)
2. Independence of hits on nonoverlapping regions
3. Equal (small) areas have equal prob. of being hit

Then there is a \( \lambda > 0 \) so that

\[
\text{#Hits in region } B
\sim \text{Poisson} \left( \mu = \lambda \cdot \text{size}(B) \right)
\]

Rain falls at rate

2 drops per square foot per second.

\[
P \left\{ \text{all 9 square feet in a square yard get hit at least once in next second} \right\}
\]

\[
= \left( 1 - P \left( N_i = 0 \right) \right)^9
= \left( 1 - e^{-2} \right)^9
= (1 - 0.1353)^9
= 0.27017
\]

(\( \lambda \)) since conditions (1), (2), (3) of Poisson Scatter Thm satisfied here.
Interpretations of Rate $\lambda$

**Global (LLN)**
\[
\lambda \approx \frac{\text{# hits in very large region } B}{\text{size } (B)}
\]

**Local**

For small region $A$, 
\[
\text{# hits} \sim \text{Pois}(\lambda_A = \frac{\lambda \text{ size } A}{\text{size } (A)})
\]
\[
P(1 \text{ hit in } A) = \lambda \text{ size } (A) e^{-\frac{\lambda \text{ size } (A)}{20}}
\]
\[
\approx \lambda \text{ size } (A)
\]

More generally, if (1), (2) and (3) Small region $A$ around point $x$ has prob. $\approx \lambda(x) \cdot \text{size}(A)$ of being hit (where $\lambda(x)$ = "intensity near $x$") then

\[
\text{# Hits in (any) region } B \sim \text{Poisson} \left( \mu_B = \int_B \lambda(x) \, dx \right)
\]
(and #Hits in disjoint sets indep)

Example of "General Poisson Scatter"

\[\text{\begin{array}{c}
\text{Sun} \\
\text{Mon} \\
\text{Tues} \\
\text{Wed} \\
\text{Th} \\
\text{Fri} \\
\text{Sat} \\
\text{Sun}
\end{array}}\]

\[\lambda(t) = \text{rate at which fatal car accidents occur at time of week } t \text{ in Tippecanoe Co.} \]
\[\#	ext{x in } [t_1, t_2] \sim \text{Pois} \left( \mu = \int_{t_1}^{t_2} \lambda(t) \, dt \right) \]
Information on Exam 1

Geometric mean $\frac{1}{p}$, variance $\frac{q}{p^2}$

For $X \sim \text{Poisson} (\mu)$,

$$P(X = k) = e^{-\mu} \frac{\mu^k}{k!}, \quad k = 0, 1, 2, \ldots$$

$$E(X) = \mu$$

$$\text{SD}(X) = \sqrt{\mu}$$

Ex Flying Bomb hits in South London

<table>
<thead>
<tr>
<th>$k$</th>
<th>$0$</th>
<th>$1$</th>
<th>$2$</th>
<th>$3$</th>
<th>$4$</th>
<th>$5+$</th>
</tr>
</thead>
<tbody>
<tr>
<td>actual #</td>
<td>229</td>
<td>211</td>
<td>93</td>
<td>35</td>
<td>7</td>
<td>1</td>
</tr>
<tr>
<td>expected $\mu$</td>
<td>226.7</td>
<td>214.4</td>
<td>98.5</td>
<td>30.6</td>
<td>7.1</td>
<td>1.6</td>
</tr>
<tr>
<td>P.S. ($\mu = 9.32$)</td>
<td>576 squares ($\frac{1}{4} \text{ km}^2$ @)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

$537$ total hits

Is Poisson Scatter with

$\mu \approx \frac{537}{576} = 0.932$ per square?

Review of Poisson Dist.

For small $p$, Bin $(n,p) \approx \text{Pois} (np)$

If $N_1 \sim \text{Poisson} (\mu_1)$, independent

$N_2 \sim \text{Poisson} (\mu_2)$

then

$N_1 + N_2 \sim \text{Poisson} (\mu_1 + \mu_2)$

Given $N_1 + N_2 = 17$,

$N_1 \sim \text{Binomial} (n=17, p=\frac{\mu_1}{\mu_1 + \mu_2})$

$X_1 + X_2 + \ldots + X_n \approx \text{Poisson} (\mu = \sum p_k)$

Independently Bernoulli ($p_k$), $p_k$ small

If Poisson scatter with

$\mu =$ mean $\#$ hits/square $= \frac{537}{576} \approx 0.932$

then

$P(0 \text{ hits in square}) = e^{-\mu}$

$= 0.3938$

$E(\# 0\text{-Hit squares})$

$$= 576 \times 0.3938$$

$$= 224.7$$

Explanation of expected $\#$'s in Flying Bomb Table

(start of Session 17)
Combining and Thinning Poisson Scatters

1. If you combine a Poisson scatter with intensity \( \lambda \) with an independent Poisson scatter with intensity \( \beta \), you get a Poisson scatter with intensity \( \lambda + \beta \).

2. If keep each point in a Poisson-\( \lambda \) scatter with prob. \( p \) and erase it with prob. \( q = 1 - p \), then
   (a) Kept points = Poisson-\( p \lambda \) scatter
   (b) Erased pts. = Poisson-\( q \lambda \) scatter
   (c) Scatters (a) and (b) independent

*Independently for diff. points.*

---

Examples

1. If rain-drops hits each second are Poisson (\( \lambda = 2 \) per square foot) scatter, then hits over 10 seconds are Poisson (\( \lambda_T = 20 \) per sq. ft.) scatter.

2. If misprints are a Poisson scatter with \( \lambda = 3 \) per page, and proofreader catches each with prob. \( \frac{2}{3} \), then remaining misprints are Poisson scatter with intensity \( \frac{1}{3} \lambda = 1 \) per page.

---

Ex. 2 (Misprints) cont.

Prob. that on 2 facing pages, 5 misprints caught, 3 uncaught?

\[ N_c = \# \text{caught} \sim \text{Poisson}(\lambda_c = 2q \lambda = 4) \]

\[ N_u = \# \text{uncaught} \sim \text{Poisson}(\lambda_u = 2p \lambda = 2) \]

\[ P\{N_c = 5, N_u = 3\} \]

\[ = \frac{e^{-\lambda_c} \lambda_c^5}{5!} \frac{e^{-\lambda_u} \lambda_u^3}{3!} \]

\[ = 0.0282 \]

---

Ex. 2, Misprints, continued again

\[ \equiv P\{N_c = 5, N_u = 3\} \]

\[ = P\{N_T = N_c + N_u = 8 , N_c = 5\} \]

\[ = \frac{e^{-6} 6^8}{8!} \left[ \frac{1}{8} \right] \left[ \frac{5}{3} \right] \left[ \frac{1}{3} \right] \]

\[ = 0.0282 \text{ as before.} \]
Law of Large Numbers
Large random samples will resemble the whole population (with high probability).

Belief in the Law of Small Numbers
Thinking that small random samples should resemble the whole population.

Problem of Multiple Testing
If you investigate 100 different types of cancer in your area, probably at least a couple of types will be found to occur at an unusually high rate.

Texas-sharpshooter Fallacy
Defining the population base around the observed cases.

Belief in the Law of Small Numbers, cont.
Tendency to see patterns in random data and to think they are due to something other than randomness.

Examples:
Residential cancer-clusters
Hot and cold streaks in B-ball
Runs of good and bad luck in gambling

Cancer cluster references:
http://bear.cba.ufl.edu/brenner/qmb 3250/clusters.htm
(Article by Gina Kolata)
(Jan-Feb 1999, Chance News)
"The Cancer-Cluster Myth"
Atul Gawande

New Yorker, Feb. 8, 1999

"The reality is that [residential cancer-cluster investigations] are an absolute, total, and complete waste of taxpayer dollars."

—Alan Bender
Minnesota Dept. of Health epidemiologist

Cancer-cluster example

McFarland, California,
pop. 6,400
11 children with cancer*
pesticides in groundwater?
Sue chemical companies!!

*but 9 different types