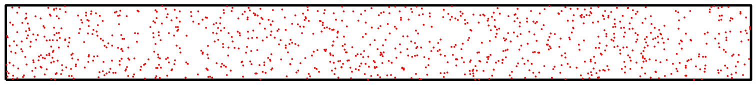
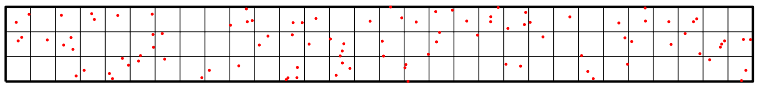
Math 180A, Fall 2005, Prof. Tesler – October 31, 2005
Poisson distribution



Collect rainfall for 1 second on a flat surface. Suppose the average density is λ (in drops per cm^2) and we are interested in a particular region of area A. What is the probability of exactly k drops in this region? Let $X = 0, 1, 2, \ldots$ be a random variable giving the number of drops in this region.



Method using the binomial distribution: Divide the region into n equal cells. If n is large enough, we can assume that the probability of two or more raindrops hitting the same cell is essentially 0. (The drawing shows $n = 90$, which apparently is not large enough.) An "event" at a cell is a raindrop hitting the cell. We also assume that the events at each cell occur independently of other cells, with equal probability in each cell. X gives the total number of cells with hits, and follows a binomial distribution with mean $\mu = \lambda A = np$ so $p = \lambda A/n = \mu/n$:

$$P(X = k) = \binom{n}{k} p^k (1-p)^{n-k} = \binom{n}{k} \left(\frac{\lambda\, A}{n}\right)^k \left(1 - \frac{\lambda\, A}{n}\right)^{n-k} = \binom{n}{k} \left(\frac{\mu}{n}\right)^k \left(1 - \frac{\mu}{n}\right)^{n-k} .$$

Suppose that in a very large area, we determine an average intensity $\lambda = 0.01\ \text{mm}^{-2}$. Then we consider a portion of that region with area $A = 123\ \text{mm}^2$. The expected number of raindrops in that area is $\mu = \lambda\, A = (.01\ \text{mm}^{-2})(123\ \text{mm}^2) = 1.23$ (a pure number with no units). What is $P(X = 3)$? We don't know what to pick for n, but as this table shows, the first several digits stabilize as n increases:

# cells n	**Probability per cell** $p = \mu/n$	**Binomial pdf at $k = 3$** $P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$
10^1	$1.23 \cdot 10^{-1}$	0.08910328876
10^2	$1.23 \cdot 10^{-2}$	0.09058485007
10^3	$1.23 \cdot 10^{-3}$	0.09064683438
10^4	$1.23 \cdot 10^{-4}$	0.09065233222
10^5	$1.23 \cdot 10^{-5}$	0.09065287510
10^6	$1.23 \cdot 10^{-6}$	0.09065292933
10^7	$1.23 \cdot 10^{-7}$	0.09065293476
10^8	$1.23 \cdot 10^{-8}$	0.09065293534

Method using the Poisson limit: The Poisson parameter is $\mu = \lambda\, A = 1.23$. Under the Poisson distribution,

$$P(X = k) = \frac{e^{-\mu} \mu^k}{k!} \qquad\qquad P(X = 3) = \frac{e^{-1.23} (1.23)^3}{3!} = 0.09065293537$$

This is a limit of the binomial distribution as used in the first method (chapter 2.4).

The probabilities in the first method converge to this value as $n \to \infty$. Even when $n = 10$, they are within 2% of each other.

Probabilities of various counts:

For a region of area $A = 123$ mm^2 and average intensity $\lambda = 0.01$ mm^{-2}, the Poisson parameter is $\mu = \lambda A = 1.23$. This table shows the probability of k events. If we look at 100 separate regions with this area, the expected number of them with exactly k events would be $100\,P(X = k)$:

# events k	**Theoretical proportion (pdf)** $P(X = k) = \dfrac{e^{-1.23}(1.23)^k}{k!}$	**Theoretical frequency** $100\,P(X = k)$
0	.2922925777	29.2292577
1	.3595198706	35.95198706
2	.2211047204	22.11047204
3	.09065293537	9.065293537
4	.02787577763	2.787577763
5	.006857441295	0.6857441295
6	.001405775465	0.1405775465
7	.0002470148317	0.02470148317
...	...	...

Determining the Poisson parameter from data. Suppose that rainfall is steady at unknown intensity λ. Do 100 separate trials in which the number of drops in a 10 mm^2 region in 1 second is measured. (It could be the same region at 100 separate times, or 100 separate regions.) Estimate the Poisson parameter and the intensity λ.

k	**Observed frequency** # trials with k events	**Observed proportion** frequency/# trials	**# events accounted for** $k \cdot$(# trials with k events)
0	64	0.64	0
1	29	0.29	29
2	6	0.06	12
3	1	0.01	3
total	100	1.00	44

The total number of events that occurred among all 100 trials is $0(64) + 1(29) + 2(6) + 3(1) = 44$ so the average number of events per trial is $44/100 = 0.44$, which we use as an estimate of the Poisson parameter μ. Since $\mu = \lambda\,A$, the average intensity is $\lambda = \mu/A = 0.44/(10 \text{ mm}^2) = 0.044$ mm^{-2} (per second). Check: $\mu = \lambda\,A = 0.44$ gives the table

k	**Theoretical proportion (pdf)** $P(X = k) = \dfrac{e^{-0.44}(0.44)^k}{k!}$	**Theoretical frequency** $100\,P(X = k)$
0	.6440364211	64.40364211
1	.2833760253	28.33760253
2	.06234272555	6.234272555
3	.009143599749	.9143599749
4	.001005795973	.1005795973
5	.00008851004558	.008851004558
...	...	...

and the entries in the "theoretical frequency" column are close to the corresponding values in the "observed frequency" column in the previous table.

For some other values not directly in this table: $P(X = 1.5) = 0$ and $P(X = -2) = 0$ (not non-negative integers); $P(X \geq 3) = 1 - P(X = 0) - P(X = 1) - P(X = 2) = .010244821$; theoretical frequency of $X \geq 3$ is $100 \cdot P(X \geq 3) = 1.0244821$.