MONASH University

Information Technology

FIT1006
Business Information Analysis

Lecture 11
Binomial and Poisson Distribution

Topics covered:

- The Binomial Distribution
- The Poisson Distribution
- Poisson approximation to Binomial
- Probability calculations (manual v calculators)

Motivating problem

- A cosmetics retailer sells 60 of a particular lipstick during the year. The shop is open 300 days per year. The shop keeps one lipstick on the counter and no others in stock. If the lipstick sells it can be replaced prior to opening the next trading day. However if a second or third... customer tries to buy the lipstick on any given day they lose the sale.
- What is the probability that the retailer loses a sale because they are out of stock?

https://flux.qa (Feed code: SJ6KGV)

Question 1 – Motivating Problem

What is the probability that the retailer loses a sale on a particular day because they are out of stock?

A. 0.002

✓ B. 0.02

C. 0.2

D. 0.5

E. None of these.

https://flux.qa (Feed code: SJ6KGV)

Question 2

Which of the following situations could not be modelled using a Binomial distribution?

A. A student guesses correct answers to a multiple choice test where each question has 4 choices.

B. Students in class today indicate whether they are left or right-handed.

✓ C. I count the number of customers who visited Chadstone shopping centre today.

D. I count the number of bad eggs in a carton of 12.

The Binomial Distribution

- We use the Binomial Distribution to determine the ‘number of successes’, each with a probability p of occurring in n independent trials.
- Typical questions:
  - A coin is tossed 50 times, what is the probability that 22 heads will be tossed? What is the probability of more than 40 heads being tossed?
  - A machine produces bolts and from past experience it is known that there is a 0.01 probability that a bolt produced will be defective. If a box contains 100 bolts, what is the probability that there are less than 5 defectives in the box.

Formal Statement

- The use of the Binomial Distribution is valid under the following conditions:
  1. Trials are independent
  2. There are only two outcomes for each trial
  3. The probability of success in each trial is constant

For n independent trials, each with a probability
p, of success, we define the the number of successes X,
as a Binomial Distribution with the following formula :

$$P(X = x) = {}^{n}C_{x}\, p^{x}(1-p)^{(n-x)} \quad for\ x = 0,1,2,3\ldots n$$

https://flux.qa (Feed code: SJ6KGV)

Question 3

Which of the following situations could not be modelled using a Binomial distribution?

A. A player tosses a coin 10 times and records the number of tails.

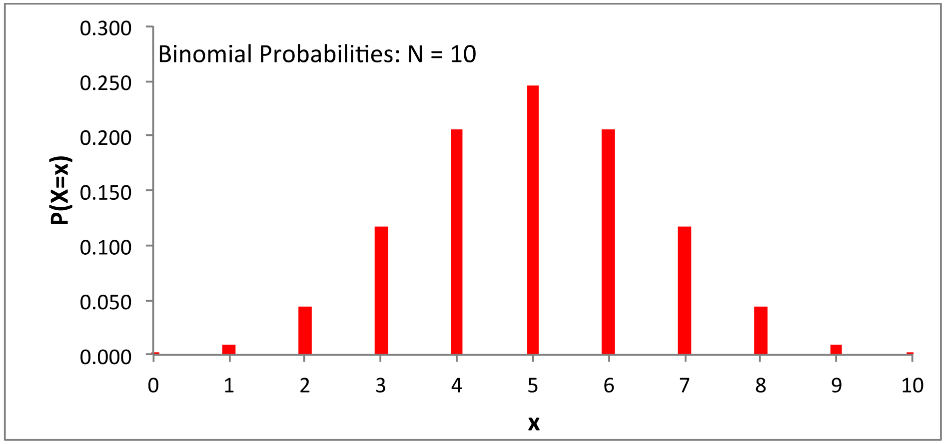
✓ B. The number of student Moodle access complaints each day is counted.

C. Throwing 6 sixes (on a die) in a row.

D. Choosing 100 people at random and recording their gender.

Example

- A fair coin is tossed 10 times and the number of heads appearing uppermost is counted. What is the probability distribution for the number of heads thrown?
- Using table constructed in Excel:
- Probabilities for each x below.



n =	**10**
p =	**0.50**
0	0.0010
1	0.0098
2	0.0439
3	0.1172
4	0.2051
5	0.2461
6	0.2051
7	0.1172
8	0.0439
9	0.0098
10	0.0010
sum =	**1.0000**

Calculating a Binomial Probability

- Example: A fair coin is tossed 12 times and the number of heads appearing uppermost is counted. What is the probability that 8 heads are thrown?

$$\text{We have } n = 12,\ p = 0.5 \text{ and } x = 8$$

$$\text{thus}$$

$$P(x) = {}^{n}C_{x}\, p^{x}(1-p)^{(n-x)}$$

$$P(8) = {}^{12}C_{8}\, 0.5^{8}(1-0.5)^{(12-8)}$$

$$= \frac{12\times 11\times 10\times 9}{4\times 3\times 2\times 1}\times 0.5^{8}\times 0.5^{4}$$

$$= 0.1208$$

Binomial Tables (n = 12)

$$\text{Recall}: \; n = 12, \; p = 0.5 \; and \; x = 8$$

Table gives P(X=x) for X = Bi(n,p)

n =	12										
p =	0.05	0.10	0.20	0.30	0.40	0.50	0.60	0.70	0.80	0.90	0.95
0	0.5404	0.2824	0.0687	0.0138	0.0022	0.0002	0.0000	0.0000	0.0000	0.0000	0.0000
1	0.3413	0.3766	0.2062	0.0712	0.0174	0.0029	0.0003	0.0000	0.0000	0.0000	0.0000
2	0.0988	0.2301	0.2835	0.1678	0.0639	0.0161	0.0025	0.0002	0.0000	0.0000	0.0000
3	0.0173	0.0852	0.2362	0.2397	0.1419	0.0537	0.0125	0.0015	0.0001	0.0000	0.0000
4	0.0021	0.0213	0.1329	0.2311	0.2128	0.1208	0.0420	0.0078	0.0005	0.0000	0.0000
5	0.0002	0.0038	0.0532	0.1585	0.2270	0.1934	0.1009	0.0291	0.0033	0.0000	0.0000
6	0.0000	0.0005	0.0155	0.0792	0.1766	0.2256	0.1766	0.0792	0.0155	0.0005	0.0000
7	0.0000	0.0000	0.0033	0.0291	0.1009	0.1934	0.2270	0.1585	0.0532	0.0038	0.0002
8	0.0000	0.0000	0.0005	0.0078	0.0420	0.1208	0.2128	0.2311	0.1329	0.0213	0.0021
9	0.0000	0.0000	0.0001	0.0015	0.0125	0.0537	0.1419	0.2397	0.2362	0.0852	0.0173
10	0.0000	0.0000	0.0000	0.0002	0.0025	0.0161	0.0639	0.1678	0.2835	0.2301	0.0988
11	0.0000	0.0000	0.0000	0.0000	0.0003	0.0029	0.0174	0.0712	0.2062	0.3766	0.3413
12	0.0000	0.0000	0.0000	0.0000	0.0000	0.0002	0.0022	0.0138	0.0687	0.2824	0.5404

Cumulative Probabilities

- Example: A fair coin is tossed 12 times and the number of heads appearing uppermost is counted. What is the probability that up to and including 8 heads are thrown?
- Use the table of cumulative probabilities!

Cumulative Binomial Table

Table gives $P(X<x)$ for $X = Bi(n,p)$

n = 12

p =	0.05	0.10	0.20	0.30	0.40	0.50	0.60	0.70	0.80	0.90	0.95
0	0.5404	0.2824	0.0687	0.0138	0.0022	0.0002	0.0000	0.0000	0.0000	0.0000	0.0000
1	0.8816	0.6590	0.2749	0.0850	0.0196	0.0032	0.0003	0.0000	0.0000	0.0000	0.0000
2	0.9804	0.8891	0.5583	0.2528	0.0834	0.0193	0.0028	0.0002	0.0000	0.0000	0.0000
3	0.9978	0.9744	0.7946	0.4925	0.2253	0.0730	0.0153	0.0017	0.0001	0.0000	0.0000
4	0.9998	0.9957	0.9274	0.7237	0.4382	0.1938	0.0573	0.0095	0.0006	0.0000	0.0000
5	1.0000	0.9995	0.9806	0.8822	0.6652	0.3872	0.1582	0.0386	0.0039	0.0001	0.0000
6	1.0000	0.9999	0.9961	0.9614	0.8418	0.6128	0.3348	0.1178	0.0194	0.0005	0.0000
7	1.0000	1.0000	0.9994	0.9905	0.9427	0.8062	0.5618	0.2763	0.0726	0.0043	0.0002
8	1.0000	1.0000	0.9999	0.9983	0.9847	0.9270	0.7747	0.5075	0.2054	0.0256	0.0022
9	1.0000	1.0000	1.0000	0.9998	0.9972	0.9807	0.9166	0.7472	0.4417	0.1109	0.0196
10	1.0000	1.0000	1.0000	1.0000	0.9997	0.9968	0.9804	0.9150	0.7251	0.3410	0.1184
11	1.0000	1.0000	1.0000	1.0000	1.0000	0.9998	0.9978	0.9862	0.9313	0.7176	0.4596
12	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

Binomial Probabilities in EXCEL

- We can use the built-in functions in EXCEL to calculate binomial probabilities.

- For $X = Bi(n,p)$

$$P(X = x) = BINOMDIST(x, n, p, false)$$

Number of successes in trials

Number of trials

Probability of success

- The cumulative probability is given by

$$P(X <= x) = BINOMDIST(x, n, p, true)$$

- … or learn to use your calculator for these.

Problem

- A farmer produces free range eggs which are collected, graded and packaged into cartons of 12. The probability that a defective egg gets through the packaging process is 0.05. What is the probability that a carton will contain more than one defective egg.
- Now, $n = 12, p = 0.05,$

  We want $P(x > 1) = P(x = 2) + P(x = 3) + \ldots P(x=12)$
- A quicker way to use our knowledge of complementary events. *
- $P(x > 1) = 1 – [P(x = 0) + P(x = 1)]$

$$= 1 - 0.5404 - 0.3413 = 0.1183$$

https://flux.qa (Feed code: SJ6KGV)

Question 4

Which of the following situations could not be modelled using a Poisson distribution?

A. The number of mutations in a strand of DNA.

✓ B. The number of damaged books in a box of 50.

C. The number of earthquakes occurring in the Pacific Rim during 2013.

D. The number of hail stones landing on a $1m^2$ area during a storm.

The Poisson Distribution

- We use the Poisson distribution to determine the number of occurrences of a random event, distributed over time or space.
- Typical Poisson Distribution questions:
  - On average 100 customers per day visit a particular shop. What is the probability that 10 customers enter the shop over one hour?
  - A fabric is known to have, on average, one defect per 10 meters. What is the probability that 5 meters of fabric will have two defects?
  - We may also use the Poisson distribution as an approximation to the Binomial distribution.

Formal Statement

- The use of the Poisson Distribution is valid under the following conditions:

1. Trials record the number of occurrences of a random event distributed over time or space.
2. The number of occurrences is theoretically unlimited

For a random event occurring with an average rate λ*,*
the mean number of occurrences over a period or area t
is given by $\mu = \lambda t$*. The number of occurences over the*
time or period x*, has a Poisson distribution with the formula:*

$$P(X = x) = \frac{e^{-\mu}\mu^{x}}{x!} \quad for\ x = 0,1,2,3\ldots\infty$$

More Poisson examples

From Wikipedia: http://en.wikipedia.org/wiki/Poisson_distribution

- Electrical system example: telephone calls arriving in a system.
- Astronomy example: photons arriving at a telescope.
- Biology example: the number of mutations on a strand of DNA per unit length.
- Management example: customers arriving at a counter or call centre.
- Civil engineering example: cars arriving at a traffic light.
- Finance and insurance example: Number of Losses/Claims occurring in a given time.
- Radioactivity Example: Decay of a radioactive nucleus.
- The number of soldiers killed by horse-kicks each year in each corps in the Prussian cavalry.
- The number of yeast cells used when brewing Guinness beer. This example was made famous by William Sealy Gosset (1876–1937).
- The number of goals in sports involving two competing teams.
- The number of deaths per year in a given age group.
- The number of jumps in a stock price in a given time interval.

Example

- A certain fabric has a probability of 0.05 that a given metre will have a defect. What is the probability that a 10m length will have 3 defects?

Per metre

10 x 0.05

$$We\ have\ t = 10,\ \lambda = 0.05\ \text{thus}\ \mu = 0.5$$

e = 2.718

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

$$P(3) = \frac{e^{-0.5} \times 0.5^{3}}{3!}$$

$$= \frac{0.6065 \times 0.125}{6}$$

$$= 0.0126$$

Poisson Tables

Table gives P(X=x) for X = Poi(mu)

mu	0.5	1.0	2.0	3.0	4.0	5.0	6.0	7.0	8.0	9.0	10.0
0	0.6065	0.3679	0.1353	0.0498	0.0183	0.0067	0.0025	0.0009	0.0003	0.0001	0.0000
1	0.3033	0.3679	0.2707	0.1494	0.0733	0.0337	0.0149	0.0064	0.0027	0.0011	0.0005
2	0.0758	0.1839	0.2707	0.2240	0.1465	0.0842	0.0446	0.0223	0.0107	0.0050	0.0023
3	0.0126	0.0613	0.1804	0.2240	0.1954	0.1404	0.0892	0.0521	0.0286	0.0150	0.0076
4	0.0016	0.0153	0.0902	0.1680	0.1954	0.1755	0.1339	0.0912	0.0573	0.0337	0.0189
5	0.0002	0.0031	0.0361	0.1008	0.1563	0.1755	0.1606	0.1277	0.0916	0.0607	0.0378
6	0.0000	0.0005	0.0120	0.0504	0.1042	0.1462	0.1606	0.1490	0.1221	0.0911	0.0631
7	0.0000	0.0001	0.0034	0.0216	0.0595	0.1044	0.1377	0.1490	0.1396	0.1171	0.0901
8	0.0000	0.0000	0.0009	0.0081	0.0298	0.0653	0.1033	0.1304	0.1396	0.1318	0.1126
9	0.0000	0.0000	0.0002	0.0027	0.0132	0.0363	0.0688	0.1014	0.1241	0.1318	0.1251
10	0.0000	0.0000	0.0000	0.0008	0.0053	0.0181	0.0413	0.0710	0.0993	0.1186	0.1251
11	0.0000	0.0000	0.0000	0.0002	0.0019	0.0082	0.0225	0.0452	0.0722	0.0970	0.1137
12	0.0000	0.0000	0.0000	0.0001	0.0006	0.0034	0.0113	0.0263	0.0481	0.0728	0.0948
13	0.0000	0.0000	0.0000	0.0000	0.0002	0.0013	0.0052	0.0142	0.0296	0.0504	0.0729
14	0.0000	0.0000	0.0000	0.0000	0.0001	0.0005	0.0022	0.0071	0.0169	0.0324	0.0521
15	0.0000	0.0000	0.0000	0.0000	0.0000	0.0002	0.0009	0.0033	0.0090	0.0194	0.0347
16	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0003	0.0014	0.0045	0.0109	0.0217
17	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0006	0.0021	0.0058	0.0128
18	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0002	0.0009	0.0029	0.0071
19	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0004	0.0014	0.0037
20	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0002	0.0006	0.0019

Cumulative Probabilities

- Customers enter a shop at an average rate of one per minute. What is the probability that in a 10 minute period more than 7 customers will enter?
- Now, λ = *1, t = 10, and mu = 10*
- We want *P(x > 7) ie P(x = 8) + P(x = 9) + …*
- *P(x > 7) = 1 - P(x = 7) - P(x = 6) - … - P(x = 0)*
- (we can use the table of cumulative probabilities here)
- *P(x > 7) = 1 - 0.2202 = 0.7798*

Cumulative Poisson Tables

Table gives P(X<=x) for X = Poi(mu)

mu	0.5	1.0	2.0	3.0	4.0	5.0	6.0	7.0	8.0	9.0	10.0
0	0.6065	0.3679	0.1353	0.0498	0.0183	0.0067	0.0025	0.0009	0.0003	0.0001	0.0000
1	0.9098	0.7358	0.4060	0.1991	0.0916	0.0404	0.0174	0.0073	0.0030	0.0012	0.0005
2	0.9856	0.9197	0.6767	0.4232	0.2381	0.1247	0.0620	0.0296	0.0138	0.0062	0.0028
3	0.9982	0.9810	0.8571	0.6472	0.4335	0.2650	0.1512	0.0818	0.0424	0.0212	0.0103
4	0.9998	0.9963	0.9473	0.8153	0.6288	0.4405	0.2851	0.1730	0.0996	0.0550	0.0293
5	1.0000	0.9994	0.9834	0.9161	0.7851	0.6160	0.4457	0.3007	0.1912	0.1157	0.0671
6	1.0000	0.9999	0.9955	0.9665	0.8893	0.7622	0.6063	0.4497	0.3134	0.2068	0.1301
7	1.0000	1.0000	0.9989	0.9881	0.9489	0.8666	0.7440	0.5987	0.4530	0.3239	0.2202
8	1.0000	1.0000	0.9998	0.9962	0.9786	0.9319	0.8472	0.7291	0.5925	0.4557	0.3328
9	1.0000	1.0000	1.0000	0.9989	0.9919	0.9682	0.9161	0.8305	0.7166	0.5874	0.4579
10	1.0000	1.0000	1.0000	0.9997	0.9972	0.9863	0.9574	0.9015	0.8159	0.7060	0.5830
11	1.0000	1.0000	1.0000	0.9999	0.9991	0.9945	0.9799	0.9467	0.8881	0.8030	0.6968
12	1.0000	1.0000	1.0000	1.0000	0.9997	0.9980	0.9912	0.9730	0.9362	0.8758	0.7916
13	1.0000	1.0000	1.0000	1.0000	0.9999	0.9993	0.9964	0.9872	0.9658	0.9261	0.8645
14	1.0000	1.0000	1.0000	1.0000	1.0000	0.9998	0.9986	0.9943	0.9827	0.9585	0.9165
15	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9995	0.9976	0.9918	0.9780	0.9513
16	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9998	0.9990	0.9963	0.9889	0.9730
17	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9996	0.9984	0.9947	0.9857
18	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9993	0.9976	0.9928
19	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9997	0.9989	0.9965
20	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9996	0.9984

Poisson Probabilities in EXCEL

- We can use the built-in functions in EXCEL to calculate Poisson probabilities.

- For $X = Poi(mu)$

$$P(X = x) = POISSON(mu, x, false)$$

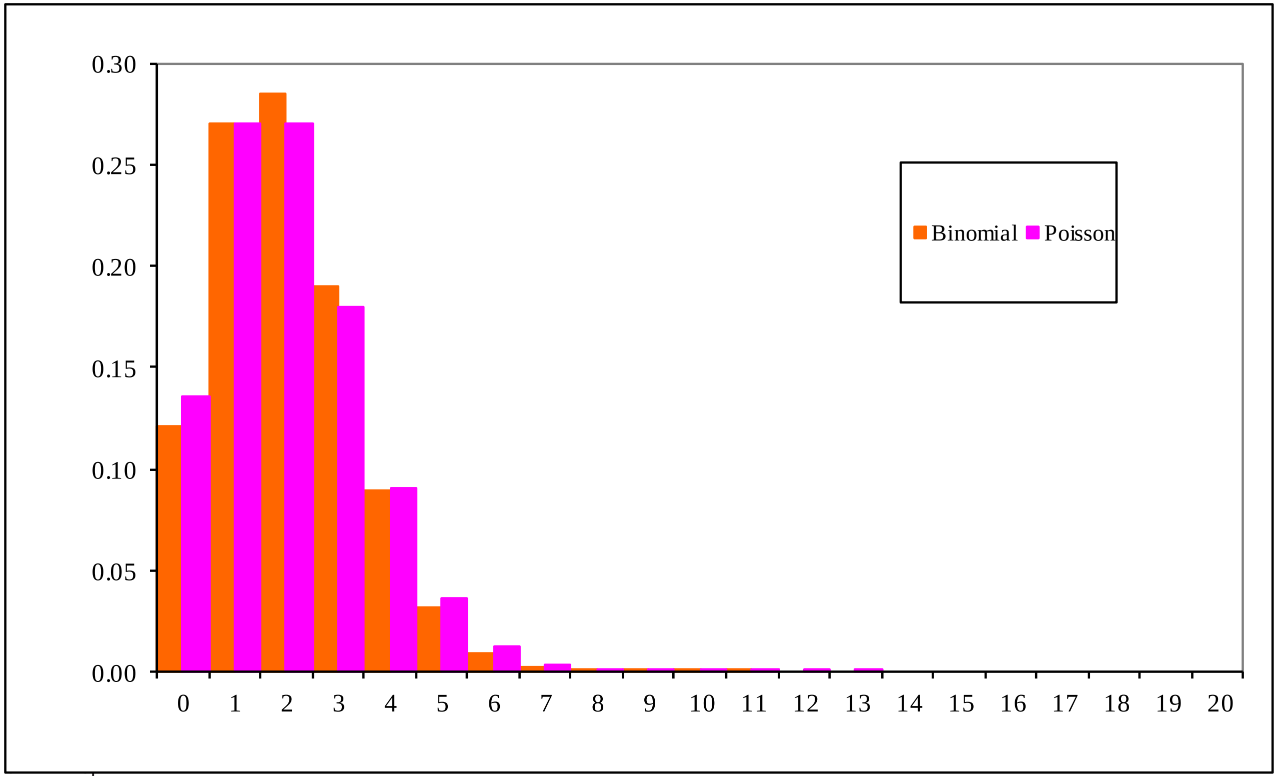
- The cumulative probability is given by

$$P(X <= x) = POISSON(mu, x, true)$$

- … or learn to use your calculator for these.

Poisson Approximation to the Binomial

- When n is large and p is small let $\mu = np$.
- The graph below shows this for $n = 20$, $p = 0.1$.



Poisson Approximation to the Binomial

- A farmer produces free range eggs which are collected, graded and packaged into cartons of 12. The probability that a defective egg gets through the packaging process is 0.05. What is the probability that a carton will contain more than one defective egg.
- Now, $n = 12, p = 0.05,$
- When n is large and p is small, we let $\mu = np$

$\mu = 12 \text{ x } 0.05 = 0.6$

- $P(X = x) = Poi(0.6)$
- $P(x > 1) = 1 - P(x = 0) - P(x = 1)$ *

$$= 1 - 0.5488 - 0.3293 = 0.1219$$

- Using a Poisson approximation to the Binomial Distribution

Sample Calculations

- For the previous question, μ = 0.05*12 = 0.6

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

$$P(0) = \frac{e^{-0.6} \times 0.6^{0}}{0!} = \frac{0.5488 \times 1}{1} = 0.5488$$

$$P(1) = \frac{e^{-0.6} \times 0.6^{1}}{1!} = \frac{0.5488 \times 0.6}{1} = 0.3293$$

$$P(2) = \frac{e^{-0.6} \times 0.6^{2}}{2!} = \frac{0.5488 \times 0.36}{2} = 0.0988$$

Mean and Variance

- The table below contains a summary of the formula for each distribution, as well as the mean and variance of each.

Distribution	Formula	Mean	Variance
Binomial	$P(X = x) = {}^{\mathrm{n}}C_x\, p^x (1-p)^{(n-x)}$	np	$np(1-p)$
Poisson	$P(X = x) = \dfrac{e^{-\mu}\mu^x}{x!}$	μ	μ

Know how to calculate Binomial and Poisson probabilities with your calculator.

Binomial *vs* Poisson

- Binomial: 2 parameters, n & p

  Number of trials

  Probability of success

  Series of n trials – Continuous

  Finite – Maximum known

  Random variable – count number of successes in fixed number of trials

- Poisson: 1 parameter, $\mu\ (= mean)$

  Discrete (0, 1, 2, …)

  Maximum theoretically infinite

  Random variable – count number of independent arrivals or independent events

Motivating problem

- A cosmetics retailer sells 60 of a particular lipstick during the year. The shop is open 300 days per year. The shop keeps one lipstick on the counter and no others in stock. If the lipstick sells it can be replaced prior to opening the next trading day. However if a second or third... customer tries to buy the lipstick on any given day they lose the sale.
- What is the probability that the retailer loses a sale because they are out of stock?

https://flux.qa (Feed code: SJ6KGV)

Question 5 – Motivating Problem

The probability that the retailer loses a sale because they are out of stock:

A. has a Binomial distribution.

✓ B. has a Poisson distribution.

Solution to the motivating problem:

$\mu = 60/300 = 0.2$

$P(x) = \dfrac{e^{-\mu}\mu^x}{x!}$	Annual demand =	60
	Days per year =	300
$P(0) = \dfrac{e^{-0.2} \times 0.5^0}{0!}$	Average daily demand =	0.2
	Distribution =	Poisson
$= 0.8187$		
Daily Demand	Probability	Lost Sale
0	0.8187	-
1	0.1637	-
2	0.0164	0.0164
3	0.0011	0.0011
4	0.0001	0.0001
5	0.0000	0.0000
Total		0.0175

Reading/Questions

- Reading:
  - 7th Ed. Sections 7.1, 7.2, 7.6, 7.7.
- Questions:
- 7th Ed. Questions 7.50, 7.53, 7.54, 7.56, 7.58, 7.60, 7.62, 7.64, 7.63, 7.68, 7.69.
- Also file: ProbabilityDistributions.xls