Probability and Statstics

Dr. Faisal Bukhari

Associate Professor

Department of Data Science

Faculty of Computing and Information Technology

University of the Punjab

Textbooks

❏ **Probability & Statistics for Engineers & Scientists**, Ninth Edition, Ronald E. Walpole, Raymond H. Myer

❏ **Elementary Statistics: Picturing the World,** 6th Edition, Ron Larson and Betsy Farber

❏ **Elementary Statistics**, 13th Edition, Mario F. Triola

Reference books

❑ **Probability Demystified**, Allan G. Bluman

❑ **Schaum's Outline of Probability and Statistics**

❑ **MATLAB Primer**, Seventh Edition

❑ **MATLAB Demystified** by McMahon, David

References

Readings for these lecture notes:

❏ Probability & Statistics for Engineers & Scientists, Ninth edition, Ronald E. Walpole, Raymond H. Myer

❏ A First Course in Probability, Eighth Edition, Sheldon Ross

❏ Elementary STATISTICS, Tenth Edition, by Mario F. Triola

References

❑ Elementary Statistics: Picturing the World, 6th Edition, Ron Larson and Betsy Farber

❑ Probability and Statistics for Computer Scientists, 2nd Edition, Michael Baron

❑ http://en.wikipedia.org/wiki/68-95-99.7_rule
These notes contain material from the above resources.

"You can make more friends in two months by becoming interested in other people than you can in two years by trying to get other people interested in you."

-Dale Carnegie

The 68-95-99.7 Rule - or Three-Sigma Rule, or Empirical Rule [1]

❑ In statistics, the **68-95-99.7 rule** or **three-sigma rule**, or **empirical rule** — states that for a normal distribution, nearly all values lie within 3 standard deviations of the mean.

❑ This rule is often used to quickly get a rough probability **estimate of something**, given its **standard deviation**, if the population is assumed normal, thus also as a simple test for **outliers** (if the population is assumed normal), and as a normality test (if the population is potentially not normal).

The 68-95-99.7 Rule - or Three-Sigma Rule, or Empirical Rule [2]

In mathematical notation, these facts can be expressed as follows, where x is an observation from a normally distributed random variable, **μ is the mean** of the distribution, and σ is its standard deviation:
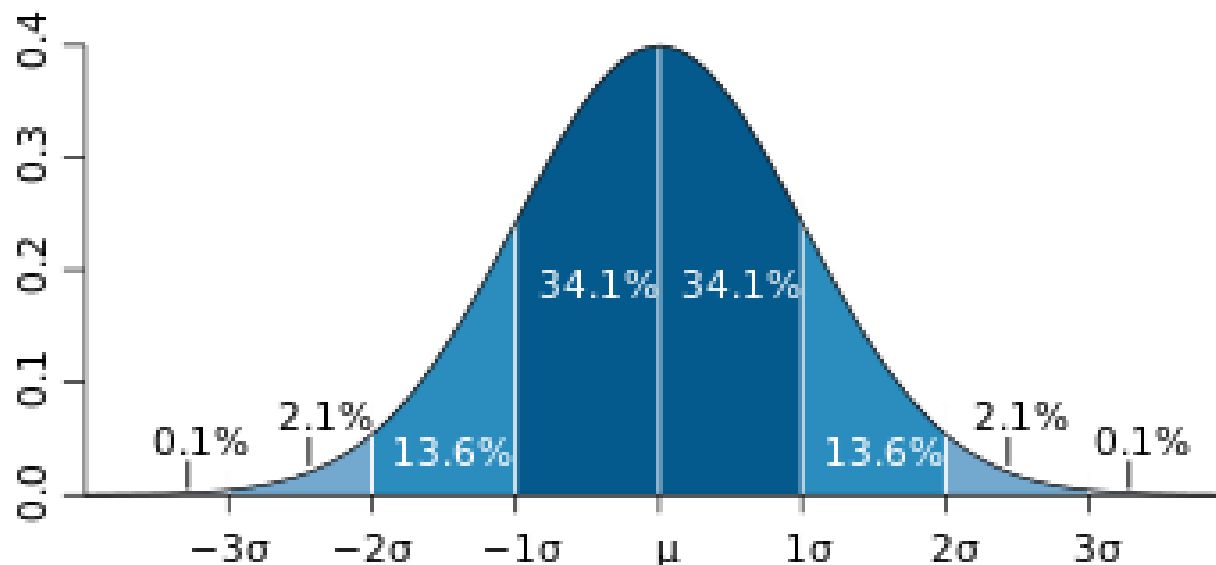
$$Pr(\mu - 1\sigma \leq x \leq \mu + 1\sigma) = \mathbf{0.6827}$$

$$Pr(\mu - 2\sigma \leq x \leq \mu + 2\sigma) = \mathbf{0.9545}$$

$$Pr(\mu - 3\sigma \leq x \leq \mu + 3\sigma) = \mathbf{0.9973}$$

The 68-95-99.7 Rule - or Three-Sigma Rule, or Empirical Rule [3]

About **68.27%** of the values lie within 1 standard deviation of the mean. Similarly, about **95.45%** of the values lie within 2 standard deviations of the mean. Nearly all **(99.73%)** of the values lie within 3 standard deviations of the mean.

Poisson Distribution

❑ In a **binomial experiment**, you are interested in **finding the probability** of a **specific number of successes** in a **given number of trials**.

❑ Suppose instead that you want to know the **probability that a specific number of occurrences** takes place within a given **unit of time, area, or volume**.

❑ For instance, to determine the probability that an employee will take 15 sick days within a year, you can use the **Poisson distribution.**

Poisson Distribution

The **Poisson distribution** is a discrete probability distribution of a random variable *x* that satisfies these conditions.

1. The experiment consists of counting the number of times *x* an event occurs in a given **interval**. The interval can be an interval of time, area, or volume. => interval is continuous

2. The probability of the event occurring is the same for each interval.

3. The number of occurrences in one interval is **independent** of the number of occurrences in other intervals.

The probability of exactly *x* occurrences in an interval is

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

where *e* = **2.71828**

Poisson Distribution [1]

Experiments yielding numerical values of a random variable *X*, the number of outcomes occurring during a given time interval or in a specified region, are called **Poisson experiments**. The given time interval may be of any length, such as a **minute**, a **day**, a **week**, a **month**, or even a **year**.

For Example, a **Poisson experiment** can generate observations for the random variable *X* representing the **number of telephone** calls received **per hour by** an office, the **number of days school** is closed due to snow during the winter, or the **number of games postponed** due to rain during a baseball season.

Properties of the Poisson Process [1]

Poisson experiment is derived from the **Poisson process** and possesses the following properties:

1. The number of outcomes occurring in one time **interval or specified region** of space is **independent** of the number that occur in any other disjoint time interval or region. In this sense we say that the **Poisson process has no memory.**

2. The probability that a **single outcome** will occur during a very **short time interval or in a small region** is proportional to the length of the time interval or the size of the region and does not depend on the number of outcomes occurring outside this time interval or region.

Properties of the Poisson Process[2]

3. The probability that **more than one outcome** will occur in such a **short time interval or fall** in such a small region is negligible.

Poisson Distribution [2]

The probability distribution of the **Poisson random variable** *X*, representing the number of outcomes occurring in a **given time interval** or specified region denoted by *t*, is

$$P(x; \lambda t) = \frac{(\lambda t)^x e^{-\lambda t}}{x!}, \quad x = 0, 1, 2, \ldots \qquad \text{λ = average not mean}$$

where **λ** is the average number of outcomes per unit time, distance, area, or volume and *e* = **2.71828**

Table A.2 contains Poisson probability sums,

$$P(r; \mu t) = \sum_{x=0}^{r} p(x; \mu t), \text{ for selected values of } \mu t$$
ranging from **0.1 to 18.0**.

Poisson Distribution [3]

❏ **Poisson distribution** is a discrete probability distribution that applies to occurrences of some event *over a **specified interval***.

❏ The random variable *x* is the number of occurrences of the event in an interval. The **interval** can be **time**, **distance**, **area**, **volume**, or some **similar unit**. The probability of the event occurring *x* times over an interval is given by

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

where *e =* **2.71828**

Poisson Distribution [4]

❏ Poisson distribution is related to a concept of **rare events**, or Poissonian events.

❏ Essentially it means that **two** such events are **extremely unlikely** to occur **simultaneously** or within a very short period of time.

❏ Arrivals of **jobs**, **telephone calls**, **e-mail messages**, **traffic accidents**, **network blackouts**, **virus attacks**, **errors in software**, **floods**, and **earthquakes** are Examples of **rare events.**

❑The number of **rare events** occurring within **a fixed period** of time has **Poisson distribution**.

Requirements for the Poisson Distribution

❑The random variable **x** is the number of occurrences of an event over **some interval**.

❑The occurrences must be **random**.

❑The occurrences must be **independent** of each other.

❑The occurrences must be **uniformly distributed** over the interval being used.

Parameters of the Poisson Distribution

❑ The mean is λt or μt

❑ The standard deviation σ is $\sqrt{\lambda t}$ or $\sqrt{\mu t}$

❑ The variance σ^2 is λt or μt

Poisson Distribution vs. Binomial Distribution

A **Poisson distribution** differs from a **binomial distribution** in these fundamental ways:

1. The **binomial distribution** is affected by the sample size n and the probability p, whereas the **Poisson distribution** is affected only by the mean.

2. In a **binomial distribution**, the possible values of the random variable x are **0, 1, . . . , n,** but a **Poisson distribution** has possible x values of **0, 1, 2, . . . ,** with **no upper limit**.

Poisson Distribution [5]

Example During a laboratory experiment, the **average** number of radioactive particles passing through a counter in **1 millisecond** is **4**. What is the probability that **6** particles enter the counter in a given **millisecond**?

Poisson Distribution [6]

Solution

Let X denotes number of particles

Here x = 6, $\mu t = (4)(1) = 4$

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

$$P(6; 4) = \frac{(4)^6 e^{-4}}{6!} = 0.1042.$$

Table A.2 Poisson Probability Sums $\sum_{x=0}^{r} p(x; \mu)$

					μ				
r	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
0	0.9048	0.8187	0.7408	0.6703	0.6065	0.5488	0.4966	0.4493	0.4066
1	0.9953	0.9825	0.9631	0.9384	0.9098	0.8781	0.8442	0.8088	0.7725
2	0.9998	0.9989	0.9964	0.9921	0.9856	0.9769	0.9659	0.9526	0.9371
3	1.0000	0.9999	0.9997	0.9992	0.9982	0.9966	0.9942	0.9909	0.9865
4		1.0000	1.0000	0.9999	0.9998	0.9996	0.9992	0.9986	0.9977
5				1.0000	1.0000	1.0000	0.9999	0.9998	0.9997
6							1.0000	1.0000	1.0000

					μ				
r	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0
0	0.3679	0.2231	0.1353	0.0821	0.0498	0.0302	0.0183	0.0111	0.0067
1	0.7358	0.5578	0.4060	0.2873	0.1991	0.1359	0.0916	0.0611	0.0404
2	0.9197	0.8088	0.6767	0.5438	0.4232	0.3208	0.2381	0.1736	0.1247
3	0.9810	0.9344	0.8571	0.7576	0.6472	0.5366	0.4335	0.3423	0.2650
4	0.9963	0.9814	0.9473	0.8912	0.8153	0.7254	0.6288	0.5321	0.4405
5	0.9994	0.9955	0.9834	0.9580	0.9161	0.8576	0.7851	0.7029	0.6160
6	0.9999	0.9991	0.9955	0.9858	0.9665	0.9347	0.8893	0.8311	0.7622
7	1.0000	0.9998	0.9989	0.9958	0.9881	0.9733	0.9489	0.9134	0.8666
8		1.0000	0.9998	0.9989	0.9962	0.9901	0.9786	0.9597	0.9319
9			1.0000	0.9997	0.9989	0.9967	0.9919	0.9829	0.9682
10				0.9999	0.9997	0.9990	0.9972	0.9933	0.9863
11				1.0000	0.9999	0.9997	0.9991	0.9976	0.9945
12					1.0000	0.9999	0.9997	0.9992	0.9980
13						1.0000	0.9999	0.9997	0.9993
14							1.0000	0.9999	0.9998
15								1.0000	0.9999
16									1.0000

Table A.2 (continued) Poisson Probability Sums $\sum_{x=0} p(x; \mu)$

	μ								
r	5.5	6.0	6.5	7.0	7.5	8.0	8.5	9.0	9.5
0	0.0041	0.0025	0.0015	0.0009	0.0006	0.0003	0.0002	0.0001	0.0001
1	0.0266	0.0174	0.0113	0.0073	0.0047	0.0030	0.0019	0.0012	0.0008
2	0.0884	0.0620	0.0430	0.0296	0.0203	0.0138	0.0093	0.0062	0.0042
3	0.2017	0.1512	0.1118	0.0818	0.0591	0.0424	0.0301	0.0212	0.0149
4	0.3575	0.2851	0.2237	0.1730	0.1321	0.0996	0.0744	0.0550	0.0403
5	0.5289	0.4457	0.3690	0.3007	0.2414	0.1912	0.1496	0.1157	0.0885
6	0.6860	0.6063	0.5265	0.4497	0.3782	0.3134	0.2562	0.2068	0.1649
7	0.8095	0.7440	0.6728	0.5987	0.5246	0.4530	0.3856	0.3239	0.2687
8	0.8944	0.8472	0.7916	0.7291	0.6620	0.5925	0.5231	0.4557	0.3918
9	0.9462	0.9161	0.8774	0.8305	0.7764	0.7166	0.6530	0.5874	0.5218
10	0.9747	0.9574	0.9332	0.9015	0.8622	0.8159	0.7634	0.7060	0.6453
11	0.9890	0.9799	0.9661	0.9467	0.9208	0.8881	0.8487	0.8030	0.7520
12	0.9955	0.9912	0.9840	0.9730	0.9573	0.9362	0.9091	0.8758	0.8364
13	0.9983	0.9964	0.9929	0.9872	0.9784	0.9658	0.9486	0.9261	0.8981
14	0.9994	0.9986	0.9970	0.9943	0.9897	0.9827	0.9726	0.9585	0.9400
15	0.9998	0.9995	0.9988	0.9976	0.9954	0.9918	0.9862	0.9780	0.9665
16	0.9999	0.9998	0.9996	0.9990	0.9980	0.9963	0.9934	0.9889	0.9823
17	1.0000	0.9999	0.9998	0.9996	0.9992	0.9984	0.9970	0.9947	0.9911
18		1.0000	0.9999	0.9999	0.9997	0.9993	0.9987	0.9976	0.9957
19			1.0000	1.0000	0.9999	0.9997	0.9995	0.9989	0.9980
20						0.9999	0.9998	0.9996	0.9991
21						1.0000	0.9999	0.9998	0.9996
22							1.0000	0.9999	0.9999
23								1.0000	0.9999
24									1.0000

					μ				
r	10.0	11.0	12.0	13.0	14.0	15.0	16.0	17.0	18.0
0	0.0000	0.0000	0.0000						
1	0.0005	0.0002	0.0001	0.0000	0.0000				
2	0.0028	0.0012	0.0005	0.0002	0.0001	0.0000	0.0000		
3	0.0103	0.0049	0.0023	0.0011	0.0005	0.0002	0.0001	0.0000	0.0000
4	0.0293	0.0151	0.0076	0.0037	0.0018	0.0009	0.0004	0.0002	0.0001
5	0.0671	0.0375	0.0203	0.0107	0.0055	0.0028	0.0014	0.0007	0.0003
6	0.1301	0.0786	0.0458	0.0259	0.0142	0.0076	0.0040	0.0021	0.0010
7	0.2202	0.1432	0.0895	0.0540	0.0316	0.0180	0.0100	0.0054	0.0029
8	0.3328	0.2320	0.1550	0.0998	0.0621	0.0374	0.0220	0.0126	0.0071
9	0.4579	0.3405	0.2424	0.1658	0.1094	0.0699	0.0433	0.0261	0.0154
10	0.5830	0.4599	0.3472	0.2517	0.1757	0.1185	0.0774	0.0491	0.0304
11	0.6968	0.5793	0.4616	0.3532	0.2600	0.1848	0.1270	0.0847	0.0549
12	0.7916	0.6887	0.5760	0.4631	0.3585	0.2676	0.1931	0.1350	0.0917
13	0.8645	0.7813	0.6815	0.5730	0.4644	0.3632	0.2745	0.2009	0.1426
14	0.9165	0.8540	0.7720	0.6751	0.5704	0.4657	0.3675	0.2808	0.2081
15	0.9513	0.9074	0.8444	0.7636	0.6694	0.5681	0.4667	0.3715	0.2867
16	0.9730	0.9441	0.8987	0.8355	0.7559	0.6641	0.5660	0.4677	0.3751
17	0.9857	0.9678	0.9370	0.8905	0.8272	0.7489	0.6593	0.5640	0.4686
18	0.9928	0.9823	0.9626	0.9302	0.8826	0.8195	0.7423	0.6550	0.5622
19	0.9965	0.9907	0.9787	0.9573	0.9235	0.8752	0.8122	0.7363	0.6509
20	0.9984	0.9953	0.9884	0.9750	0.9521	0.9170	0.8682	0.8055	0.7307
21	0.9993	0.9977	0.9939	0.9859	0.9712	0.9469	0.9108	0.8615	0.7991
22	0.9997	0.9990	0.9970	0.9924	0.9833	0.9673	0.9418	0.9047	0.8551
23	0.9999	0.9995	0.9985	0.9960	0.9907	0.9805	0.9633	0.9367	0.8989
24	1.0000	0.9998	0.9993	0.9980	0.9950	0.9888	0.9777	0.9594	0.9317
25		0.9999	0.9997	0.9990	0.9974	0.9938	0.9869	0.9748	0.9554
26		1.0000	0.9999	0.9995	0.9987	0.9967	0.9925	0.9848	0.9718
27			0.9999	0.9998	0.9994	0.9983	0.9959	0.9912	0.9827
28			1.0000	0.9999	0.9997	0.9991	0.9978	0.9950	0.9897
29				1.0000	0.9999	0.9996	0.9989	0.9973	0.9941
30					0.9999	0.9998	0.9994	0.9986	0.9967
31					1.0000	0.9999	0.9997	0.9993	0.9982
32						1.0000	0.9999	0.9996	0.9990
33							0.9999	0.9998	0.9995
34							1.0000	0.9999	0.9998
35								1.0000	0.9999
36									0.9999
37									1.0000

Poisson Distribution [7]

Example (alternative approach): During a laboratory experiment, the average number of radioactive particles passing through a counter in **1 millisecond** is 4. What is the probability that **6 particles** enter the counter in a **given millisecond**?

Poisson Distribution [6]

Solution:

Let X denotes number of particles

Here x = 6, μt = (4)(1) = 4

Using **Table A.2**, we get

$$P(6; 4) = \sum_{x=0}^{x=6} p(x; 4) - \sum_{x=0}^{x=5} p(x; 4) = 0.8893 - 0.7851$$
$$= 0.1042.$$

Poisson Distribution [7]

Example Ten is the **average number** of oil tankers arriving each day at a certain port. The facilities at the port can handle at most **15 tankers per day**. What is the probability that on a given day tankers have to be turned away?

Poisson Distribution [8]

Solution:

Here μ = 10, t = 1 day, and μt = 10

Let **X denotes number of tankers**

P(X > 15) = 1 − P(X ≤ 15)

$$= 1 - \sum_{x=0}^{x=15} p(x; 10)$$

$$= 1 - 0.9513 = 0.0487.$$

Poisson Distribution [9]

Example 4: Ten is the **average number** of **oil tankers** arriving each day at a certain port. The facilities at the port can handle **at most 15 tankers per day**. What is the probability that on a given day tankers have to be **turned away**?

Poisson Distribution [10]

Solution(alternative approach):

Here μ = 10, t = 1 day, and μt = 10

Let X denotes number of tankers

P(X > 15) = 1 − P(X ≤ 15)

$\quad\quad$ = 1 − [P(X = 0) + P(X = 1) + P(X = 2) + P(X = 3) + P(X = 4) + P(X = 5) + P(X = 6) + P(X = 7) + P(X = 8) + P(X = 9) + P(X = 10)+P(X = 11) + P(X = 12) + P(X = 13) + P(X = 14) + P(X = 15)]

$\quad\quad$ = 1 − 0.9513 = 0.0487.

Poisson Distribution [11]

Example Customers of an internet service provider initiate new accounts at the average rate of **5 accounts per day**.

(a) What is the probability that **more than 8 new accounts** will be **initiated today**?

(b) What is the probability that **more than 16 accounts** will be initiated within **2 days**?

(a) Here μ = 5, **t = 1 day**, and **λt = 5**

Let **X** denotes number of **new accounts**

P(X > 8) = 1 − P(X ≤ 8)

$$= 1 - \sum_{x=0}^{x=8} p(x;\ 5)$$

$$= 1 - 0.9319$$

$$= \textbf{0.0681}$$

(b) Here λ = 5, **t = 2 day**, and **λt = 10**

P(X >16) = 1 − P(X ≤ 16)

$$= 1 - \sum_{x=0}^{x=16} p(x;\ 10)$$

$$= 1 - 0.9730$$

$$= \textbf{0.027}$$

Mean and Variance of the Poisson Distribution

Both the mean and the variance of the Poisson distribution $P(x; \lambda t)$ are **λt** or $P(x; \mu t)$ are **μt**

Mean = λt or **μt**

Variance = λt or **μt**

Nature of the Poisson Probability Function

Like so many **discrete** and **continuous distributions**, the form of the **Poisson distribution** becomes more and more symmetric, even bell-shaped, as the **mean grows** large.
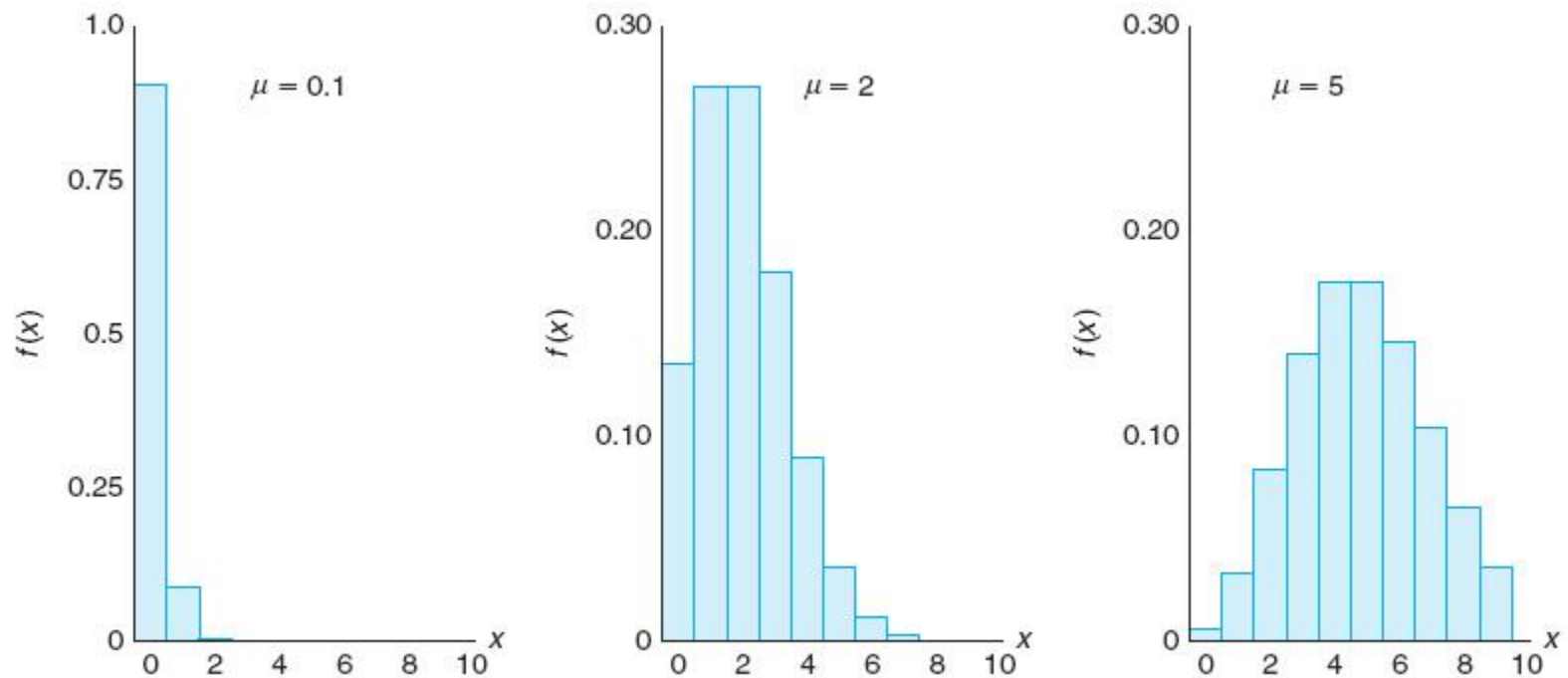
Poisson as Approximation to Binomial

The **Poisson distribution** is sometimes used to **approximate** the **binomial distribution** when *n* **is large** and *p* **is small**.

One rule of thumb is to use such an approximation when the following two conditions are both satisfied

Requirements for Using the Poisson Distribution as an approximation to the Binomial

1. $n \geq 100$

2. $np \leq 10$

According to two rules of thumb, this approximation is good if **n ≥ 20 and p ≤ 0.05, or if n ≥ 100 and np ≤ 10**

If both of these conditions are satisfied and we want to use the Poisson distribution as an approximation to the binomial distribution, we need a value for and that value can be calculated by using Formula **$\mu = np$**

Approximation of Binomial Distribution by a Poisson Distribution [1]

It should be evident from the three principles of the Poisson process that the **Poisson distribution** is related to the **binomial distribution**.

Although the **Poisson usually** finds applications in **space** and **time problems**, it can be viewed as a limiting form of the **binomial distribution.**

Approximation of Binomial Distribution by a Poisson Distribution [2]

If *n* **is large** and *p* **is close to 0**, the Poisson distribution can be used, with $\boldsymbol{\mu = np}$, to **approximate binomial** probabilities.

Let *X* be a binomial random variable with probability distribution $b(x; n, p)$. When $n \longrightarrow \infty, p \longrightarrow 0,$ **and** $np \longrightarrow \mu$ remains constant, $b(x; n, p) \longrightarrow p(x; \mu).$

❑ The **Poisson random variable** has a wide range of applications in a variety of areas because it may be used as an **approximation** for a **binomial random variable with parameters (n, p)** when **n is large and p is small**.

❑ In other words, if **n independent trials**, each of which results in a **"success"** with **probability p**, are performed, then **when n is large** and **p small**, the number of successes occurring is approximately a Poisson random variable with **mean λ = np**.

Applications

❑ Some examples of random variables that usually obey, to a good approximation, the Poisson probability law are:

1. The number of misprints on a page (or a group of pages) of a book.

2. The number of people in a community living to 100 years of age.

3. The number of wrong telephone numbers that are dialed in a day.

4. The **number of transistors** that fail on their **first day** of use.

5. The number of customers entering a post office on a **given day**.

6. The number of α-particles discharged in a fixed **period of time** from some radioactive particle.

Approximation of Binomial Distribution by a Poisson Distribution [3]

Example In a certain industrial facility, accidents occur infrequently. It is known that the probability of an accident on any given day is **0.005** and accidents are independent of each other.

(a) What is the probability that in any given period of **400** days there will be an accident on **one day**?

(b) What is the probability that there are at most **three days** with an **accident**?

Solution

Let X be a binomial random variable with **$n = 400$** and

$p = 0.005$. Thus, **$np = 2, \mu = 2$**.

According to two rules of thumb, this approximation is good if **$n \geq 20$ and $p \leq 0.05$, or if $n \geq 100$ and $np \leq 10$**

(a) $P(X = 1) = e^{-2}2^{1} = 0.271$

(b) $P(X \leq 3) = \sum_{0}^{3} \dfrac{e^{-2}2^{x}}{x!} = 0.857$

Approximation of Binomial Distribution by a Poisson Distribution [4]

Example: In a manufacturing process where glass products are made, defects or bubbles occur, occasionally rendering the piece undesirable for marketing. It is known that, on average, **1** in every **1000** of these items produced has one or more bubbles.

What is the probability that a random sample of **8000** will yield **fewer than 7** items possessing bubbles?

SOLUTION: This is essentially a binomial experiment with n = 8000 and p = $\frac{1}{1000}$ = 0.001.

μ = np

μ = (8000)(0.001) = 8.

According to two rules of thumb, this approximation is good if **n ≥ 20 and p ≤ 0.05, or if n ≥ 100 and np ≤ 10**

Let X represents the **number of glasses possessing bubbles**

$P(X < 7) = \sum_{x=0}^{x=6} b(x; 8000, 0.001)$

$\approx \sum_{x=0}^{x=6} p(x; 8) = 0.3134.$

Example Suppose the probability that an item produced by a certain machine will be **defective is 0.1**. Find the probability that a **sample of 10 items** will contain at **most one defective** item. Assume that the quality of successive items is independent.

Solution Method 1

Here n = 10   **(Total number of items)**

p = 0.10   **(Probability of defective)**

q = 1 − 0.10 = 0.90   **(Probability of non-defective)**

Let X denotes number of defective items

$$b(x; n, p) = \binom{n}{x} p^x q^{n-x}, \quad x = 0, 1, 2, ..., n$$

$$\therefore P(X \leq 1) = \binom{10}{0}(0.10)^0 (0.90)^{10} + \binom{10}{1}(0.10)^1(0.90)^9 = \textbf{0.7361}$$

Method 2

Here n = 10　　　　　　　　**(Total number of items)**

p = 0.10　　　　　　　　　**(Probability of defective)**

q = 1 − 0.10 = 0.90　　　**(Probability of non-defective)**

According to two rules of thumb, this approximation is good if **n ≥ 20 and p ≤ 0.05, or if n ≥ 100 and np ≤ 10**

Let X denotes number of defective items

λ = np = (10)(0.10) = 1

$$P(x; 1) = \frac{(1)^x e^{-1}}{x!} \quad , x = 0, 1, 2, \ldots$$

$$\therefore \boldsymbol{P(X \leq 1) = \frac{(1)^0 e^{-1}}{0!} + \frac{(1)^1 e^{-1}}{1!}} = 2e^{-1} \approx \textbf{0.7358}$$

Poisson Distribution

Example: An automobile manufacturer is concerned about a fault in the braking mechanism of a particular model. The fault can, on rare occasions, cause a catastrophe at high speed. The distribution of the number of **cars per year** that will experience the fault is a Poisson random variable with $\lambda = 5$.

(a) What is the probability that **at most 3** cars per year will experience a catastrophe?

(b) What is the probability that **more than 1** car per year will experience a catastrophe?

Solution: Here $\lambda t=(5)(1)=5$

$$P(x; \lambda t) = \frac{(\lambda t)^x e^{-\lambda t}}{x!}, x = 0, 1, 2, . . .$$

(a) P(X ≤ 3) $= \sum_{x=0}^{x=3} p(x; 5)$
$$= 0.2650$$

(b) P(X > 1) $= 1 - P(x \leq 1)$
$$= 1 - \sum_{x=0}^{x=1} p(x; 5)$$
$$= 1 - 0.0404$$
$$= 0.9596$$

Poisson Distribution

Example: Changes in airport procedures require considerable planning. Arrival rates of aircraft are important factors that must be taken into account. Suppose small aircraft arrive at a certain airport, according to a Poisson process, at the rate of **6 per hour**. Thus the Poisson parameter for arrivals for a period of hours is **$\lambda = 6$**.

(a) What is the probability that **exactly 4** small aircraft arrive during a **1-hour period**?

Poisson Distribution

(b) What is the probability that **at least 4** arrive during a **1-hour period**?

(c) If we define a **working day** as **12 hours**, what is the probability that at least **75 small aircraft** arrive **during a day**?

Poisson Distribution

Here $\lambda t = (6)(1) = 6$

$$P(x; \lambda t) = \frac{(\lambda t)^x e^{-\lambda t}}{x!}, \ x = 0, 1, 2, \ldots$$

(a) $P(X = 4) = \dfrac{(6)^4 e^{-6}}{4!} = $ **0.1339**

(b) $P(X \geq 4) = 1 - P(x < 4) = 1 - \sum_{x=0}^{x=3} p(x; 6) = 1 - 0.1512 = $ **0.8488**

Here **$\lambda t = (6)(12) = 72$**

(c) $P(X \geq 75) = 1 - P(x < 75) = 1 - \sum_{x=0}^{x=74} p(x; 72) = $ **0.3773.**

Poisson Distribution using MATLAB

poisspdf is Poisson probability density function in Matlab.

Y = poisspdf(X, LAMBDA) returns the Poisson probability density function with parameter LAMBDA at the values in X

Poisson Distribution using MATLAB

a)

```
x = 4
reqprob = poisspdf(x, 6)
% 0.1339
```

b)

```
x = 0:3
prob = poisspdf(x, 6)
Reqprob = 1 - sum(prob)
%0.8488
```

Poisson Distribution using MATLAB

c)

```
%  vector notation
x = 0:74
% sum of probability

prob = sum(poisspdf(x, 72))

% required probability
reqprob = 1 – prob

%  0.3773
```

Poisson approximation

The **Binomial distribution** converges towards the **Poisson distribution** as the number of trials goes to **infinity** while the product **np** remains fixed. Therefore the Poisson distribution with parameter **λ= np** can be used as an approximation to b(n, p) of the binomial distribution if n is sufficiently large and p is sufficiently small.

According to two rules of thumb, this approximation is good if

n ≥ 20 and p ≤ 0.05, or if n ≥ 100 and np ≤ 10.

Poisson Distribution [4]

Formula:

$$P(x; \lambda) = \frac{e^{-\lambda} \lambda^x}{x!}, \; x = 0, 1, 2, \ldots$$

where, λ is an average rate of value, x is a Poisson random variable and e is the base of logarithm(e = 2.718).

Example:

Consider, in an office on **average 2 customers** arrived per day. Calculate the possibilities for exactly 3 customers to be arrived on today.

Step1: Find $e^{-\lambda t}$.
where, $\lambda t = 2$ and $e = 2.718$, $e^{-\lambda t} = (2.718)^{-2} = 0.135$.

Step2: Find $(\lambda t)^x$
where, $t = 1$, $\lambda t = 2$ and $x = 3$, $(\lambda t)^x = 2^3 = 8$.

Step3: Find $P(x; \lambda)$
$$P(x; \lambda) = \frac{e^{-\lambda t}(\lambda t)^x}{x!}, \; x = 0, 1, 2, \ldots$$

$$P(3; 2) = \frac{(0.135)(8)}{3!} = 0.18.$$
Hence there are 18% possibilities for 3 customers to be arrived today

Geometric Distribution

❑ Many actions in life are **repeated until** a success occurs.

For instance, you might have to send an **e-mail several** times before it is **successfully sent**. A situation such as this can be represented by a **geometric distribution.**