Probability and Statistics MAT 271E

PART 5

Bernoulli Trial and Poisson Process

Assist. Prof. Dr. Ümit KARADOĞAN

Course originally developed by : Prof. Dr. Mehmetçik BAYAZIT, Prof. Dr. Beyhan YEĞEN

BERNOULLI TRIALS

Let us consider an experiment where only **two outcomes** are possible (There are only two simple events in the sample space.)

Suppose one of the outcomes corresponds (arbitrarily) to "**success**" and the other to "**failure**".

The probability of the **success** will be denoted by ***p*,**
and the probability of the **failure** by ***q=1-p*.**

Such an experiment is called a ***<u>Bernoulli trial</u>*.**

BERNOULLI TRIALS

As a simple example, if the "**success**" in the throw of a die is equated with the throw of a **six**, then $p=1/6$ and $q=5/6$.

Let us repeat the experiment **n times** (**independent Bernoulli trials**)

Now consider the random variable **X**, the **number of times of success** in ***n*** trials.

X is a **discrete** variable, that is an integer in the range of 0 to *n.*

Let us compute the probability of $X=\boldsymbol{x}$ in $\boldsymbol{n}$ trials.

BERNOULLI TRIALS

Suppose **n=3 (three trials)**.

The event of **no success** *(X=0)* will occur only when all the three trials are failures.

Since the trials are considered independent, this has the probability:

$P(X=0) = q\ q\ q = q^3$

1 success in 3 trials can occur in three different ways:

first trial successful and the others failures

second trial successful and the others failures

third trial successful and the others failures

Each of these three events has the probability $p\ q^2$ and the probability of their union is:

$P(X=1) = 3\ pq^2$

BERNOULLI TRIALS

2 successes in 3 trials can also occur in three different ways:
first two trials successful and the third trial failure
first and third trials successful and the second failure
second and third trials successful and the first failure
Each event has the probability p^2q and therefore the probability of 2 successes in 3 events is:

$P(X=2) = 3\,p^2q$

Finally the probability of **3 successes in 3 events** is:

$P(X=3) = p^3$

BERNOULLI TRIALS

Generalizing to **n trials**, the probability of **x successes** can be computed as

$$P(X=x)=\binom{n}{x}p^x q^{n-x} \qquad x=0,1,\ldots,n$$

where $\binom{n}{x}$ denotes the number of **combinations** of ***n* different things *taken x at a time***, to be computed as:

$$\binom{n}{x}=\frac{n!}{x!\,(n-x)!}$$

The distribution of the variable X is called the **Bernoulli distribution (termed as Binomial distribution for n trials)**.

BERNOULLI TRIALS

The parameters of the variable *X* are:

The expected number of successes in **n** trials is, as expected, equal to **np**.

E(X) =np   Var(X)=npq

For 1 trial:

$$\begin{aligned} \mathrm{E}[X] &= \sum_{x\in\{0,1\}} x\, p^x (1-p)^{1-x} \\ &= 0\cdot p^0(1-p)^{1-0} + 1\cdot p^1(1-p)^{1-1} \\ &= 0 + p \\ &= p \end{aligned}$$

$$\begin{aligned} \mathrm{E}[X^2] &= \sum_{x\in\{0,1\}} x^2\, p^x (1-p)^{1-x} \\ &= 0^2\cdot p^0(1-p)^{1-0} + 1^2\cdot p^1(1-p)^{1-1} \\ &= 0 + p \\ &= p. \end{aligned}$$

$$\begin{aligned} \mathrm{Var}[X] &= \mathrm{E}[X^2] - (\mathrm{E}[X])^2 \\ &= p - p^2 \\ &= p(1-p) \end{aligned}$$

Example (M. Bayazıt, B. Oğuz, Example 2.23, pg 52)

The probability of a **successful** bid for a contractor is assumed to be *p=1/3.*
Let us compute the probability of 0,1,2, and 3 successes in 3 bids.
The variable ***X*** denotes the **number of successes** in 3 bids.

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

$\boxed{\frac{n!}{x!(n-x)!}\text{=3!/(1!x2!)}}$

P(X=0)=$\binom{3}{0}(1/3)^0(2/3)^3$=8/27
P(X=1)=$\binom{3}{1}(1/3)^1(2/3)^2$ =12/27
P(X=2)=$\binom{3}{2}(1/3)^2(2/3)^1$ =6/27
P(X=3)=$\binom{3}{3}(1/3)^3(2/3)^0$ =1/27

Example (M. Bayazıt, B. Oğuz, Example 2.23, pg 52)

The probability that the contractor is successful **at least once in 3** bids can be computed as follows:

$P[X \geq 1]=1-P[X<1]=1-P[X=0]=1-8/27=19/27$

The expected value of successes in 3 bids is:

$E[X] = =np = 3 \times 1/3 =1$

Example (M. Bayazıt, B. Oğuz, Example 2.23, pg 52)

We can compute the probability of the **first success** to occur in the **y-th trial** as follows.

This will happen when the **first y-1 trials are failures**, and the next trial is a success.

The probability of y-1 failures is q^{y-1} and the probability of success at the y-th trial is p. Therefore:

$$P(Y=y)=q^{y-1}p \qquad y=1,2...$$

Geometric Distribution

This is called the ***Geometric distribution***. Its parameters are:

$E(Y)=1/p$ $\qquad$ $Var(Y)=q/p^2$

As expected, a success will occur on average, once in every **p** trials.

In the example of the throw of a die:

the probability of a **six in the first trial** is $P(Y=1)=1/6$

in the **second trial** $P(Y=2)=(5/6)(1/6) = 5/36$, etc.

We can expect to throw a six once in every six trials on average.

Example (M. Bayazıt, B. Oğuz, Example 2.24, pg 53)

In the previous example, the probabilities of first **success** in the **first, second** and **third**, ... bids are:

$P(Y=1)=(2/3)^0(1/3)=1/3$

$P(Y=2)=(2/3)(1/3)=2/9$

$P(Y=3)=(2/3)^2(1/3)=4/27$

...

Return period (recurrence interval) T is defined as the average interval between two consecutive successes.

This coincides with the **average time to the first success**.

$$T=1/p$$

Example

The lifetime in hours of a certain kind of radio tube is a random variable having a probability density function given by:

$$f(x)=\begin{cases} 0 & x \leq 100 \\ \dfrac{100}{x^2} & x > 100 \end{cases}$$

What is the probability that exactly 2 of 5 such tubes in a radio set will have to be replaced within the first 150 hours operation?

Assume that the events that the i-th tube will have to be replaced are independent.

Example

$$\mathrm{P}(E_i) = \int_0^{150} f(x)dx$$

$$= 100\int_{100}^{150} x^{-2} dx$$

$$= \frac{1}{3}$$

$$\binom{5}{2}(1/3)^2(2/3)^3 = \frac{80}{243}$$

Example (M. Bayazıt, B. Oğuz, Example 2.25, pg 53)

The spillway of a dam is designed for a discharge that is exceeded with the probability of 0.01 each year.

The life of the dam is assumed to be 50 years.

What is the probability that the design flow is exceeded (at least once) during the life of the project?

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

Using **Bernoulli distribution**:

$P(X\geq1)=1-P(X=0)=1-\binom{50}{0}(0.01)^0(0.99)^{50}=0.395$

What is the probability that the first exceedance occurs after more than 10 years?

Using the equation for the **geometric distribution**: $\mathbf{P(Y=y)=q^{y-1}p}$ **y= 1, 2,...**

$$P(Y>10)=1-\sum_{y=1}^{10}P(Y=y)=1-\sum_{y=1}^{10}0.99^{y-1}0.01=0.92$$

Example (M. Bayazıt, B. Oğuz, Example 2.25, pg 53)

What is the average interval between two consecutive exceedances of flood?

$E(Y)=1/p$

$E(Y)= 1/0.01= 100$ years

T=1/p is also the **return period** of the design flood

T= 100 years

Which means that the design flood (or a larger flood) will occur, on the average, once every 100 years (this is called the 100-year flood: a flood that has 1 in 100 chance (1% probability) of being equaled or exceeded in any given year.)

It should not be understood that this event (exceedance of the design flood) will occur at intervals of 100 years.

In fact, there is a probability of $0.01^2=0.0001$ that it will occur twice in two consecutive years.

Example (M. Bayazıt, B. Oğuz, Example 2.25, pg 53)

If the spillway was designed for the 50-year flood (p=1/50), the probability of no such flood occuring during the life of the project would be:

$$P(Y=0)=\binom{50}{0}(1/50)^{0}(49/50)^{50}=0.364$$

Example (M. Bayazıt, B. Oğuz, Example 2.25, pg 53)

The probability of **no occurrence** of an event with the return period of T years during a time interval of T years can be computed as:

$P(Y=0)=(1-p)^T=1-Tp+\frac{T(T-1)}{2}p^2+\ ...$

which approaches e^{-Tp} for large values of T.

Therefore:

$P(Y=0)=e^{-Tp} = e^{-\frac{1}{p}p} = e^{-1} = 0.368$

The value of 0.364 obtained for T=50 years is very close to this.

Example (M. Bayazıt, B. Oğuz, Example 2.30, pg 56)

The probability of the strength of a construction material to fall below the standards is given as p=0.20.

What is the probability that **at most 2** elements have strengths below the standards in a sample of n=30 elements?

Bernoulli distribution with n=30, x=2:

$$P(X \leq 2) = \sum_{x=0}^{2} \binom{30}{x} (0.20)^{x} (0.80)^{30-x} = 0.044 \cong 0.05$$

If we accept that the material conforms to the standards when **not more than 2** elements in the sample have strengths below the standards, then we allow a risk of about 0.05.

Example (M. Bayazıt, B. Oğuz, Example 2.30, pg 56)

On the other hand, if we accept the material when **not more than 1** element in the sample has strength below the standards, the the risk will be:

$$P(X \le 1) = \sum_{x=0}^{1} \binom{30}{x} (0.20)^x (0.80)^{30-x} \cong 0.01$$

Accepting the material only when **all the elements** in the sample have the required strength will reduce the risk further:

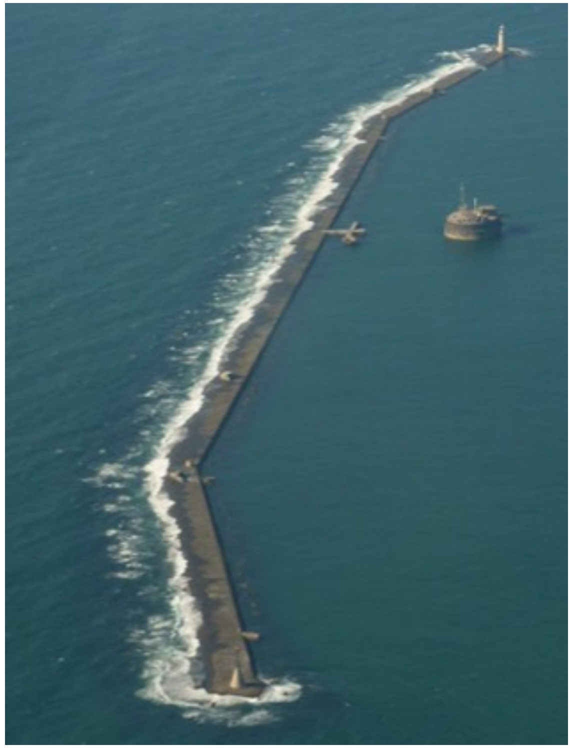
$$P(X=0) = 0.80^{30} \cong 0.001$$

Example (M. Bayazıt, B. Oğuz, Example 2.26, pg 54)

A breakwater is designed for waves of the return period T=10 years (probability of exceedance p=0.10). The probability that the breakwater is damaged when larger waves occur is assumed to be 0.20.

What is the probability of damage in a 3 year period?

The probability of larger waves occuring X=0, 1, 2, and 3 years are calculated by :



$P(X=0)=\binom{3}{0}(0.10)^0(1-0.10)^3=0.729$

$P(X=1)=\binom{3}{1}(0.10)^1(1-0.10)^2=0.243$

$P(X=2)=\binom{3}{2}(0.10)^2(1-0.10)^1=0.027$

$P(X=3)=\binom{3}{3}(0.10)^3(1-0.10)^0=0.001$

Example (M. Bayazıt, B. Oğuz, Example 2.26, pg 54)

Using the total probability theorem, the probability of no damage in a 3 year period is computed as:

$1.0 \times 0.729 + (1\text{-}0.20)\ 0.243 + (1\text{-}0.20)^2\ 0.027 + (1\text{-}0.20)^3\ 0.001 = 0.94$

The probability of damage in 3 years is:

$1\text{-}0.94 = 0.06$

If the number of the independent Bernoulli trials is increased ($n \to \infty$), at the same time decreasing the probability of success in a trial ($p \to 0$), in such a way that the average number of successes in n trials is kept constant ($np = \nu$), then the probability of x successes occuring in n trials will approach to the following expression.

Example (M. Bayazıt, B. Oğuz, Example 2.26, pg 54)

$$P(X=x) = \frac{\nu^x e^{-\nu}}{x!} \qquad x=0, 1, 2, \ldots$$

where $\nu=E(X)$.

This limit form of the Bernoulli distribution is called the **Poisson distribution**.

Its parameters are:

$$E(X) = Var(X) = \nu$$

The Poisson distribution is useful in computing the probabilities P(X=x) for large number of trials n.

POISSON PROCESS

The **Poisson** distribution is useful for calculating:
the **number of times an event occurs in an interval of time or space**.

Assume that the **probability of a random event to occur** in a **small time interval** **Δt** is proportional to the **length of Δt**, and that the consecutive occurrences of the event are **independent**.

The **distribution of x (the number of occurrences)** in the interval (0, t), can be determined using the **Poisson distribution**.

Expected value of the number of occurrences is:

$\boxed{E(X) = \nu = \lambda t \quad \textbf{(mean)}}$

Using Poisson equation: $\boxed{P(X=x) = \frac{\nu^{x} e^{-\nu}}{x!}}$

The probability mass function: $\boxed{P(X=x) = \frac{(\lambda t)^{x} e^{-\lambda t}}{x!}}$ x=0, 1, 2,...

The process defined by the above equation is called the **Poisson process**.

POISSON PROCESS

Let us determine the probability distribution of the variable T, **the time to the first occurrence** of the event for a **Poisson process**.

No occurrence of the event in the time interval (0, t) corresponds to T>t. Therefore:

$$P(T > t) = P(X=0)$$

$F(t) = 1 – P(T > t) = 1 – P(X=0)$ occurrence of the event in the time interval t

$$= 1 - \frac{(\lambda t)^0 e^{-\lambda t}}{0!} = \boxed{1 - e^{-\lambda t}}$$

Example (M. Bayazıt, B. Oğuz, Example 2.27, pg 55)

Let us compute the probability of the design flow to be exceeded in 50 years for the spillway (of Example 2.25) using **Poisson distribution**.

$\nu = np = 50 \times 0.01 = 0.5$

$P(X=0) = \dfrac{0.5^0 e^{-0.5}}{0!} = e^{-0.5} = 0.607$

$P(X=x) = \dfrac{\nu^x e^{-\nu}}{x!}$

$P(X\geq 1) = 1 - 0.607 = 0.393$ (at least once)

which is very close to the exact value of 0.395.

Example (M. Bayazıt, B. Oğuz, Example 2.28, pg 55)

In an intersection, the average number of accidents in a week is 3 (ν = 3).
What is the probability of 5 accidents in a week ?

Poisson distribution with x=5:

P(X=5) = $\frac{3^5 e^{-3}}{5!}$ = 0.10

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

Example (M. Bayazıt, B. Oğuz, Example 2.29, pg 55)

It is assumed that the number of vehicles waiting at red light at an intersection obeys the **Poisson distribution** with a mean of ν =10.

The intersection is blocked when more than 15 cars are waiting.

What is the probability of this event?

Using Poisson distribution with x=15:

$$P(X>15) = 1 - P(X \leq 15) = 1 - \sum_{x=0}^{15} \frac{10^x e^{-10}}{x!} = 0.0487$$

Example (M. Bayazıt, B. Oğuz, Example 2.31, pg 57)

The mean number of accidents at an intersection in t=7 days is 3.

What is the probability of no accidents in t =2 days ?

$$\nu = 7\lambda = 3 \qquad \lambda = 3/7$$

For t=2 $\quad F(t) = \dfrac{(\lambda t)^0 e^{-\lambda t}}{0!} = 0.424$

$$P(X = x) = \frac{(\lambda t)^x e^{-\lambda t}}{x!}$$

The probability of no occurance in 2 days is 0.424.

Similarly, the probability of X=1 accident in 2 days is:

$$P(X = 1) = \left(\frac{3}{7} \times 2\right)^1 e^{-\frac{3}{7}\times 2} / 1! = 0.364$$

Example (M. Bayazıt, B. Oğuz, Example 2.32, pg 57)

16 earthquakes of the magnitude exceeding 6 have been observed at a location in 125 years.

What is the probability that such an earthquake will occur in the next 2 years?

Assuming that the occurence of earthquakes follow the **Poisson distribution**, the mean number of earthquakes in one year is:

$$\lambda = 16/125 = 0.128$$

The probability that the time to the first earthquake is less than or equal to 2 years is:

$$F(t) = 1 - P(X=0)$$

$$= 1 - \frac{(\lambda t)^0 e^{-\lambda t}}{0!} = 1 - e^{-\lambda t}$$

$$P(T \leq 2) = 1 - e^{-0.128x2} = 0.226$$

Example (M. Bayazıt, B. Oğuz, Example 2.32, pg 57)

The same result can be obtained using Bernoulli distribution.

The probability of no earthquakes (X=0) in 2 years is:

$$P(X=0)=\binom{2}{0}(0.128)^0(1-0.128)^2=0.774$$

The probability of an earthquake in 2 years is:

1 - 0.774 = 0.226

Example The dust particles in the atmosphere is an environmental problem.
The number of dust particles in a unit volume is measured by taking 100 measurements.
The Poisson distribution is valid for the situation.

Number of dust in unit volume	0	1	2	3	4	5
Frequency observed	13	24	30	18	7	8

The mean for dust:
E(X)= 13/100x 0 + 24/100x 1+ 30/100x 2 + 18/100 x 3 + 7/100 x 4 + 8/100 x 5

$$\bar{x} = \upsilon = 2.14$$

$$P(X = x) = \frac{(\lambda t)^x e^{-\lambda t}}{x!}$$

$$P[X = x] = \frac{2.14^x e^{-2.14}}{x!} \quad \rightarrow \quad x = 0, 1, 2, 3, 4, 5$$

Example

Number of dust in unit volume	0	1	2	3	4	5
Frequency observed	13	24	30	18	7	8
Frequency using Poisson	12	26	27	19	10	6