Seoul National University
Dept. of Civil and Environmental Engineering

Instructor: Junho Song
junhosong@snu.ac.kr

457.212 Statistics for Civil & Environmental Engineers

In-Class Material: Class 12

Useful Distribution Models (A&T: 3.2)

<< Bernoulli Trials >>

- Each trial has only (   ) possible **outcomes**
  ~ 'success'/'failure', 'occur'/'do not occur'
- Probability of occurrence of the event in each trial is (   ).
- The probabilities in different trials are **statistically** (   ).
  e.g. tossing coins repeatedly.
  occurrence of flooding each year.

1. **Binomial** distribution

(a) X ~ Binomial (n, p)

- **number of occurrences** of an event of specified probability p during n Bernoulli trials.

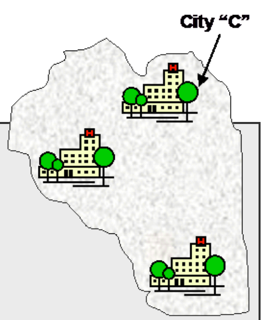
(b) PMF

$$p_X(x) = \binom{n}{x} \cdot p^x (1-p)^{n-x}$$

where $\binom{n}{x}$ is the binomial coefficient, i.e. $\dfrac{n!}{(n-x)!x!}$

(c) Mean: $\mathrm{E}[X] = np$

(d) Variance: $\sigma_X^2 = np(1-p)$



Example 1: Three hospitals in City "C." The probability that each hospital will experience the shortage of electricity is 0.1 and they are statistically independent. X is the number of hospital(s) that has electricity shortage.

$P(X = 0) =$

$P(X = 1) =$

$P(X = 2) =$

$P(X = 3) =$

X is a Binomial random variable with $n =$, $p =$.

PMF:

Mean, $\mu_X =$, and standard deviation, $\sigma_X =$

Instructor: Junho Song
junhosong@snu.ac.kr

2. **Geometric** distribution

(a) X ~ Geometric (p)

- number of trials **until the** (   ) or (   ) occurrence of an event of specified probability p

(b) PMF: $p_X(x) = (1-p)^{x-1}p$ because (x–1) no occurrences and then one occurrence

(c) Mean: $\mu_X = \mathrm{E}[X] = \sum_{x=1}^{\infty} x \cdot (1-p)^{x-1} p = \dfrac{1}{p}$

→ **Return period**: average number of trials until the first (next) occurrence

(d) Variance: $\sigma_X^2 = (1-p)/p^2$

3. **Negative Binomial** distribution

(a) X ~ Negative Binomial (k, p)

- number of trials **until the** (   ) occurrence of an event of specified probability p

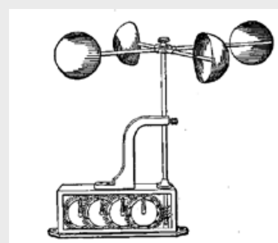
(b) PMF

$$p_X(x) = \binom{x-1}{k-1} p^{k-1}(1-p)^{x-k} \times p = \binom{x-1}{k-1} p^k (1-p)^{x-k}$$

(c) Mean: $\mu_X = \mathrm{E}[X] = \sum_{x=1}^{\infty} x \cdot p_X(x) = \dfrac{k}{p}$

(d) Variance: $\sigma_X^2 = k \cdot (1-p)/p^2$

Example 2: Event A: the annual maximum wind velocity is greater than 100 mph. P(A) = 0.2. The outcomes of all different years are statistically independent of each other.



(a) Bernoulli trials? Why? p = ?

(b) Probability that A occurs twice for a five-year duration?

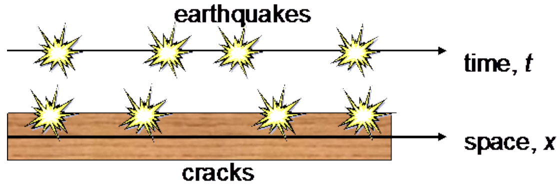
(c) Probability that A occurs for the first time at the 7th year?

(d) Return period?

(e) Probability that it will take 10 years until the 3rd occurrence?

<< Poisson Process >>

- An event can occur **at random** at any time or **any point** in space. (limiting case of Bernoulli trials as $n \to \infty$)
- The occurrence of an event in a given time (or space) interval is **statistically** (   ) of that in any other non-overlapping interval.
- The probability of occurrences of an event in a small interval Δt is proportional to Δt, i.e. **Prob.** = $\nu \Delta t$ where ν is the **constant** mean occurrence rate, i.e. (average # of occurrences)/(length)



4. **Poisson** distribution

(a) $X \sim$ Poisson (ν)

- **number of the (   ) ** of an event in time duration t for a Poisson process with ν.

(b) PMF

$$p_X(x) = \frac{(\nu t)^x}{x!} \exp(-\nu t), \quad x = 0,1,...$$

(c) Mean: $\mu_X = \mathrm{E}[X] = \sum_{x=0}^{\infty} x \cdot p_X(x) = \nu \cdot t$

(d) Variance: $\sigma_X^2 = \nu \cdot t$

Example 3: Assume the occurrences of "heavy" rainstorms follow a Poisson process. From historical data, four "heavy" rainstorms occur per year in average.

(a) The mean occurrence rate?

(b) Probability of three storms for two years?

(c) Probability of at least one rainstorm for two years?

Seoul National University
Dept. of Civil and Environmental Engineering

Instructor: Junho Song
junhosong@snu.ac.kr

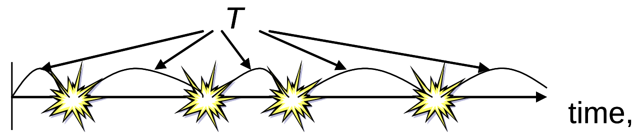
5. **Exponential** distribution

(a) T ~ Exponential (ν)

- **waiting time until the (   ) or (   ) occurrence** in a Poisson process with ν.

(b) CDF

$$F_T(t) = P(T \le t)$$
$$= P(\text{at least one occurrence in } t)$$
$$= 1 - \frac{(\nu t)^0}{0!}\exp(-\nu t)$$
$$= 1 - \exp(-\nu t)$$

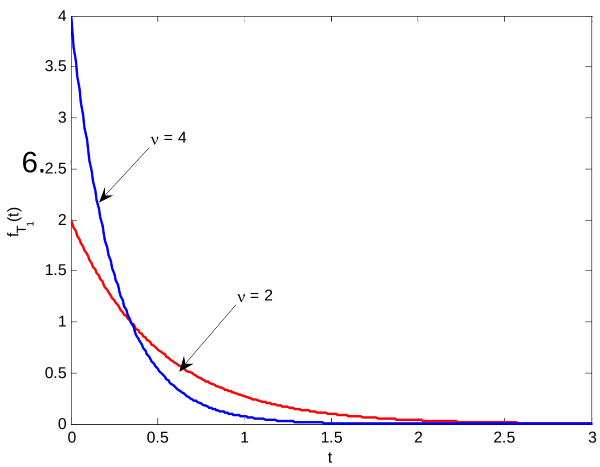


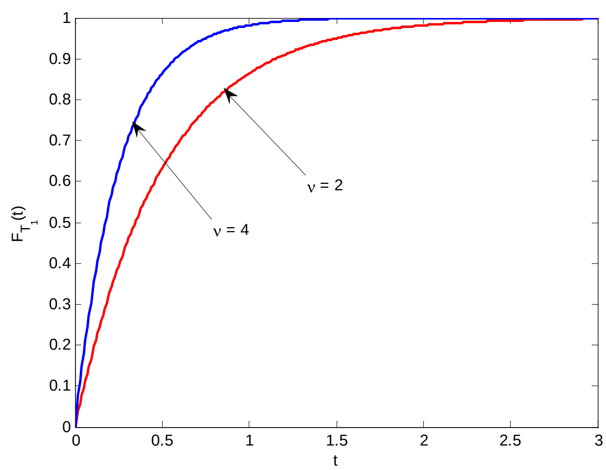
(c) PDF

$$f_T(t) = \frac{dF_T(t)}{dt}$$
$$= \nu \exp(-\nu t)$$

(c) Mean: $\mu_T = \mathrm{E}[T] = \int_0^\infty t \cdot \nu \exp(-\nu t) dt = \frac{1}{\nu}$ → **Return Period** (c.f. 1/*p* for Bernoulli)

(d) Variance: $\sigma_T^2 = \frac{1}{\nu^2}$





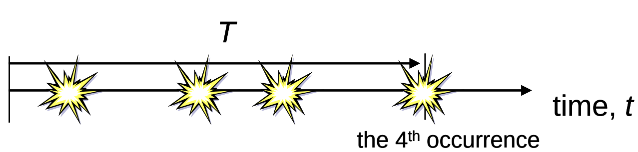
6. **Gamma** distribution

(a) T ~ Gamma (ν, *k*)

- **waiting time until the (   ) occurrence** in a Poisson process with ν.

(b) CDF:

$$F_T(t) = 1 - \sum_{x=0}^{k-1} \frac{(\nu t)^x}{x!}\exp(-\nu t)$$

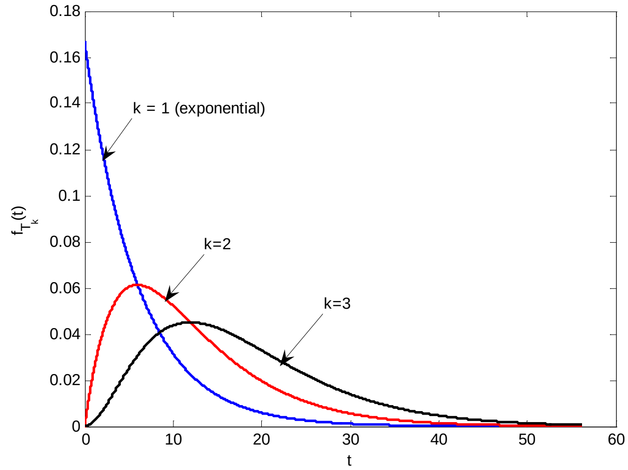


(c) PDF:

$$f_T(t) = \frac{\nu(\nu t)^{k-1}}{\Gamma(k)} \exp(-\nu t)$$

(d) Mean: $\mu_T = \mathrm{E}[T] = \frac{k}{\nu}$

(d) Variance: $\sigma_T^2 = \frac{k}{\nu^2}$



<< Summary >>

	Bernoulli Trials	Poisson Process
No. of **total trials** (or time intervals Δt)	*n* trials, or until the *k*-th occurrence	$n \to \infty$, $\Delta t \to 0$ any point on a continuous axis (time, space)
P(**no. of occurrence** = *x*)	**Binomial** (*n*, *p*) during *n* trials	**Poisson** (ν) during time period *t*
P(no. of trials or waiting time **until the first** (next) occurrence = *t*)	**Geometric** (*p*)	**Exponential** (ν)
P(no. of trials or waiting time **until the *k*-th** occurrence = *t*)	**Negative Binomial** (*k*, *p*)	**Gamma** (*k*, ν)
Return period: average required no. of trials or waiting time until the next occurrence	1/*p* where *p* is the probability of the event	1/ν where ν is the mean occurrence rate

7. **Beta** distribution

(a) X ~ Beta (*a*, *b*, *q*, *r*)

A random variable whose values are **bounded** between finite limits *a* and *b*. Its shape is determined by the combination of *q* and *r*.

(b) PDF

$$f_X(x) = \frac{1}{B(q,r)} \frac{(x-a)^{q-1}(b-x)^{r-1}}{(b-a)^{q+r-1}}, \quad a \le x \le b$$

$$= 0 \qquad \text{elsewhere}$$

where $B(q,r) = \int_0^1 x^{q-1}(1-x)^{r-1}dx = \frac{\Gamma(q)\Gamma(r)}{\Gamma(q+r)}$