SOME IMPORTANT PROBABILITY DISTRIBUTIONS

Discrete Probability Distributions

Aim - you will be able to

- characterize Bernoulli trials and types of discrete distributions
- characterize Poisson process and Poisson distribution
- describe context between discrete distributions

Bernoulli trials:

- a sequence of Bernoulli trials is defined as a sequence of random events which are mutually independent and which have only two possible outcomes (e.g. success-nonsuccess, 1-0)
- probability of event occurrence (a success) p is constant in any trial

$$\mathrm{P}\{\text{Trial 'i'} = \text{"Success"}\} = p$$

Binomial random variable:

$X \rightarrow B(n, p)$

X is the number of successes in n Bernoulli trials where the probability of success at each trial is p

$$P(X = k) = \binom{n}{k} p^k (1-p)^{n-k};\ 0 \le k \le n,$$ see below

Example:

$$P[\,S\ S\ F\ S\ F\,] = p^3\,(1-p)^2$$

$$P[\,S\ F\ F\ S\ S\,] = p^3\,(1-p)^2$$

$$P[\,F\ S\ S\ F\ S\,] = p^3\,(1-p)^2$$

$$P[\overbrace{S\quad S\quad F\quad S\quad F\quad \ldots\quad \ldots\quad F}^{\text{k successes \& n-k failures}}] = p^k\,(1-p)^{n-k}$$

$$P[X = k] = \binom{n}{k} p^k (1-p)^{n-k};\ 0 \le k \le n$$

$$EX = \sum_{k=0}^{n} k \cdot P(X = k) = \sum_{k=0}^{n} k \cdot \frac{n!}{(n-k)!k!} \cdot p^k (1-p)^{n-k} = n \cdot p \sum_{k=0}^{n} \frac{(n-1)!}{(n-k)! \cdot (k-1)!} \cdot p^{k-1} \cdot (1-p)^{n-k} = n \cdot p$$

$$DX = EX^2 - (EX)^2$$

$$EX^2 = \sum_{k=1}^{n} k^2 \cdot P(X = k) = \sum_{k=1}^{n} k \cdot (k-1) \cdot \frac{n!}{(n-k)! \cdot k!} \cdot p^k \cdot (1-p)^{n-k} + EX =$$

$$= n \cdot (n-1) \cdot p^2 \cdot \sum_{k=2}^{n} \frac{(n-2)!}{(n-k)! \cdot (k-2)!} \cdot p^{k-2} \cdot (1-p)^{n-k} + EX =$$

$$= n \cdot (n-1) \cdot p^2 + n \cdot p = (np)^2 - np^2 + np$$

Binomial theorem (secondary school): $(\alpha+\beta)^n = \sum_{k=0}^{n} \binom{n}{k} \alpha^k \beta^{n-k}$

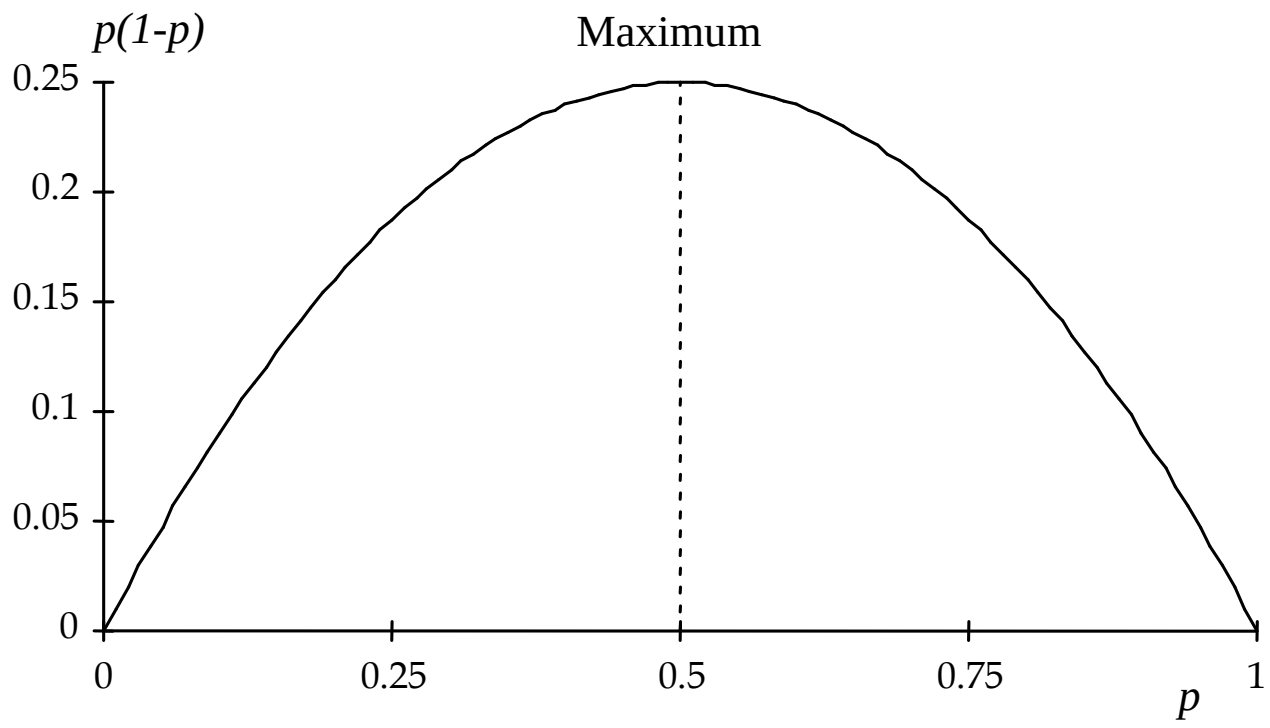
For α=1, β=1 we obtain $\sum_{k=0}^{n} \binom{n}{k} = \sum_{k=0}^{n} \binom{n}{k} 1^k 1^{n-k} = (1+1)^n = 2^n$

for α=p, β=1-p we obtain $\sum_{k=0}^{n} P[X=k] = \sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} = [p+(1-p)]^n = 1$

i.e. the Binomial RV is **normed**.

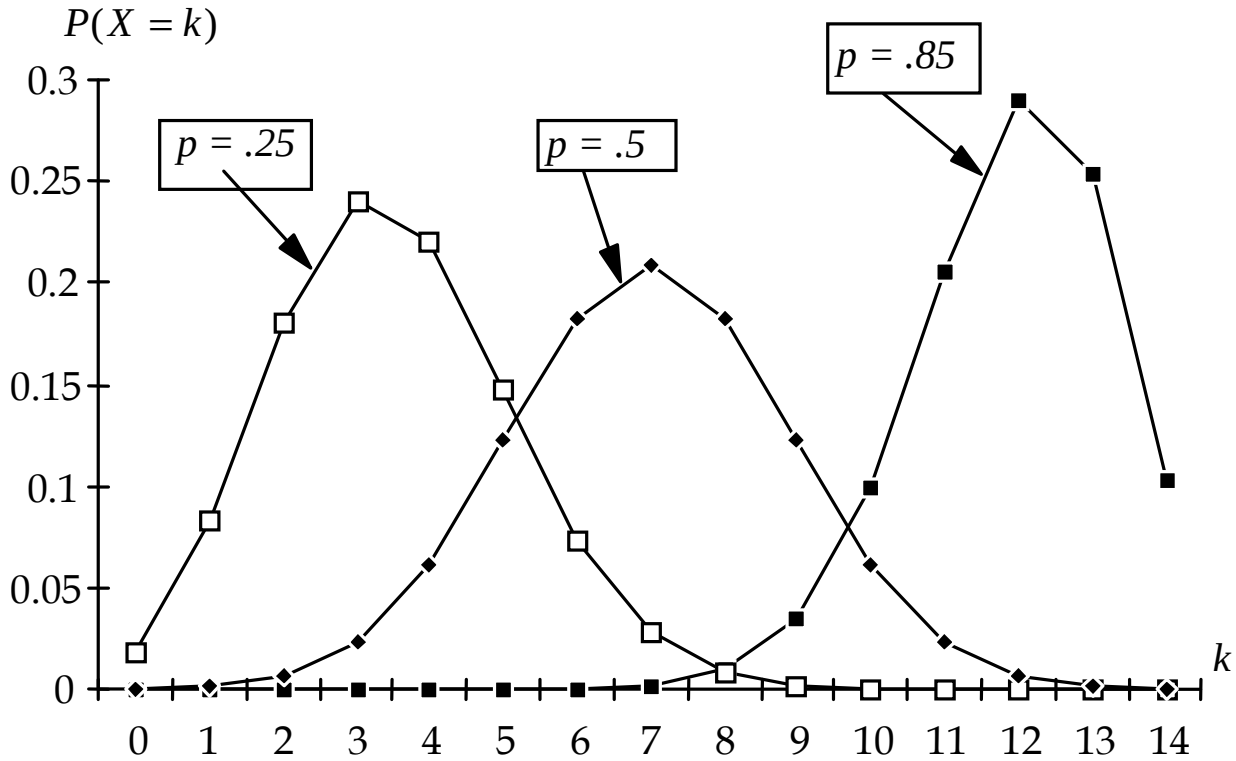
$DX = EX^2 - (EX)^2 = n.p.(1 - p)$

Notice that the variance of the binomial distribution is maximum when $p = 0.5$.



Example:

Some examples of binomial distributions for $n = 14$ trials are illustrated below. Notice that as p, the probability of success at each trial increases, the location of the distribution shifts to higher values of the random variable. Also notice that when p = 0.5, the distribution is symmetric around 7.5.



Geometric distribution:

$X \rightarrow G(p)$

G(p) … *G(p)* … the geometric random variable is defined as the number of Bernoulli trials until the first success occurs

$$P(X = k) = p\,(1-p)^{k-1};\ 1 \le k < \infty$$

$$EX = \sum_{k=1}^{\infty} k \cdot P(X = k) = \sum_{k=1}^{\infty} k \cdot p \cdot (1-p)^{k-1} = p \cdot \sum_{k=1}^{\infty} k \cdot (1-p)^{k-1} = p \cdot \frac{\partial}{\partial(1-p)} \cdot \left(\sum_{k=1}^{\infty} (1-p)^{k} \right) = \frac{1}{p}$$

Variance of $G(p)$

$$DX = EX^2 - (EX)^2 = \frac{1-p}{p^2}$$

$$E(X^2) = \sum_{k=1}^{\infty} k^2 \; p(1-p)^{k-1}$$

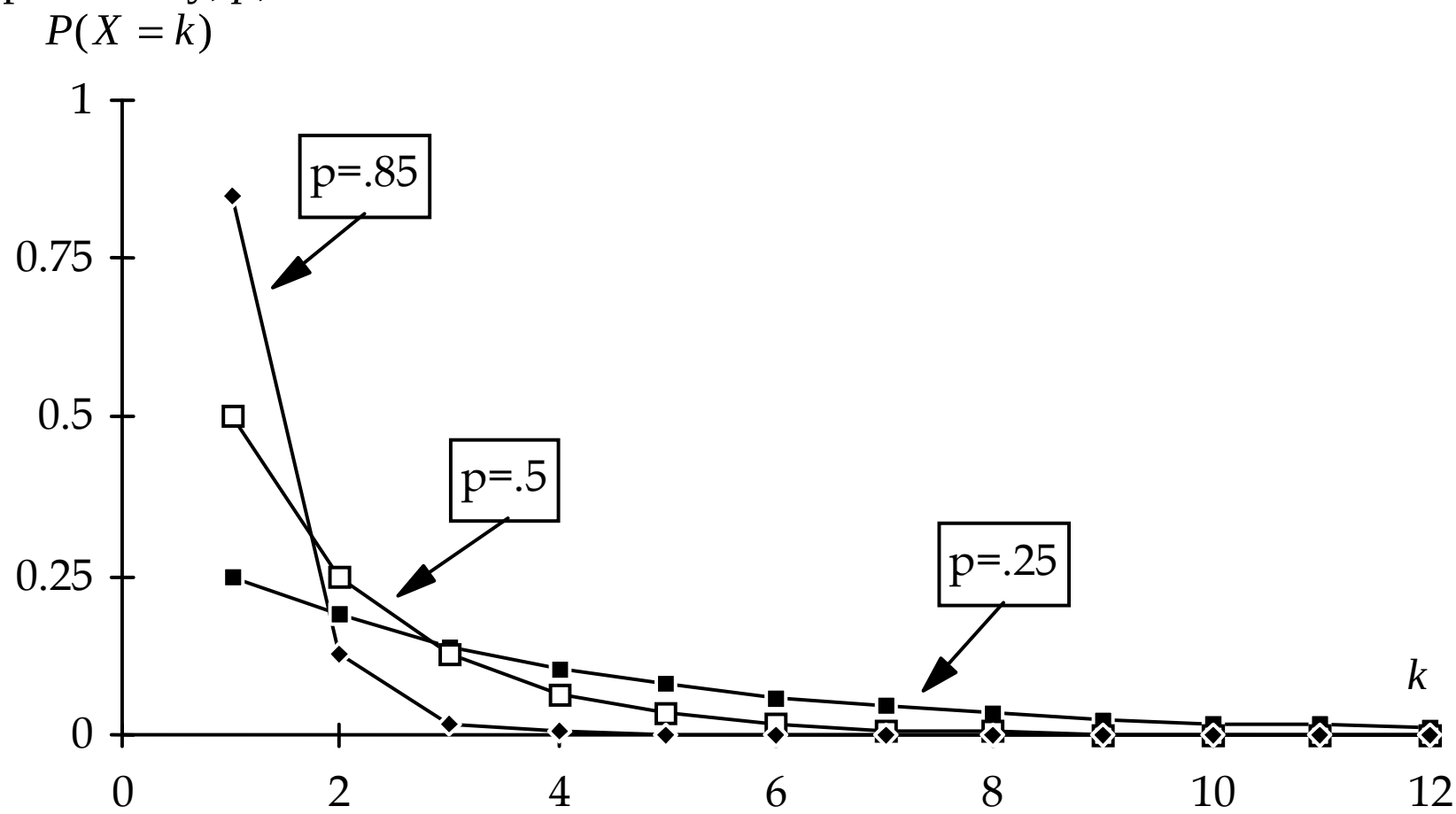
$$= p(1-p)\sum_{k=2}^{\infty} k(k-1)(1-p)^{k-2} + p\sum_{k=2}^{\infty} k(1-p)^{k-1}$$

$$= p(1-p)\frac{\partial^2 \sum_{k=1}^{\infty}(1-p)^k}{\partial^2(1-p)} + p\frac{\partial \sum_{k=1}^{\infty}(1-p)^k}{\partial(1-p)}$$

$$= p(1-p)\frac{\partial^2\left(\frac{1-p}{p}\right)}{\partial^2(1-p)} + p\frac{\partial\left(\frac{1-p}{p}\right)}{\partial(1-p)} = \frac{2(1-p)}{p^2} + \frac{1}{p}$$

Example:

Some examples of geometric distributions are illustrated below. Not surprisingly, the probability of long sequences without success decreases rapidly as the success probability, p, increases.



Negative binomial random variable

The negative binomial distribution has two parameters, k and p and is denoted by

$$X \rightarrow NB(k, p)$$

The negative binomial is the number of Bernoulli trials until the kth success.

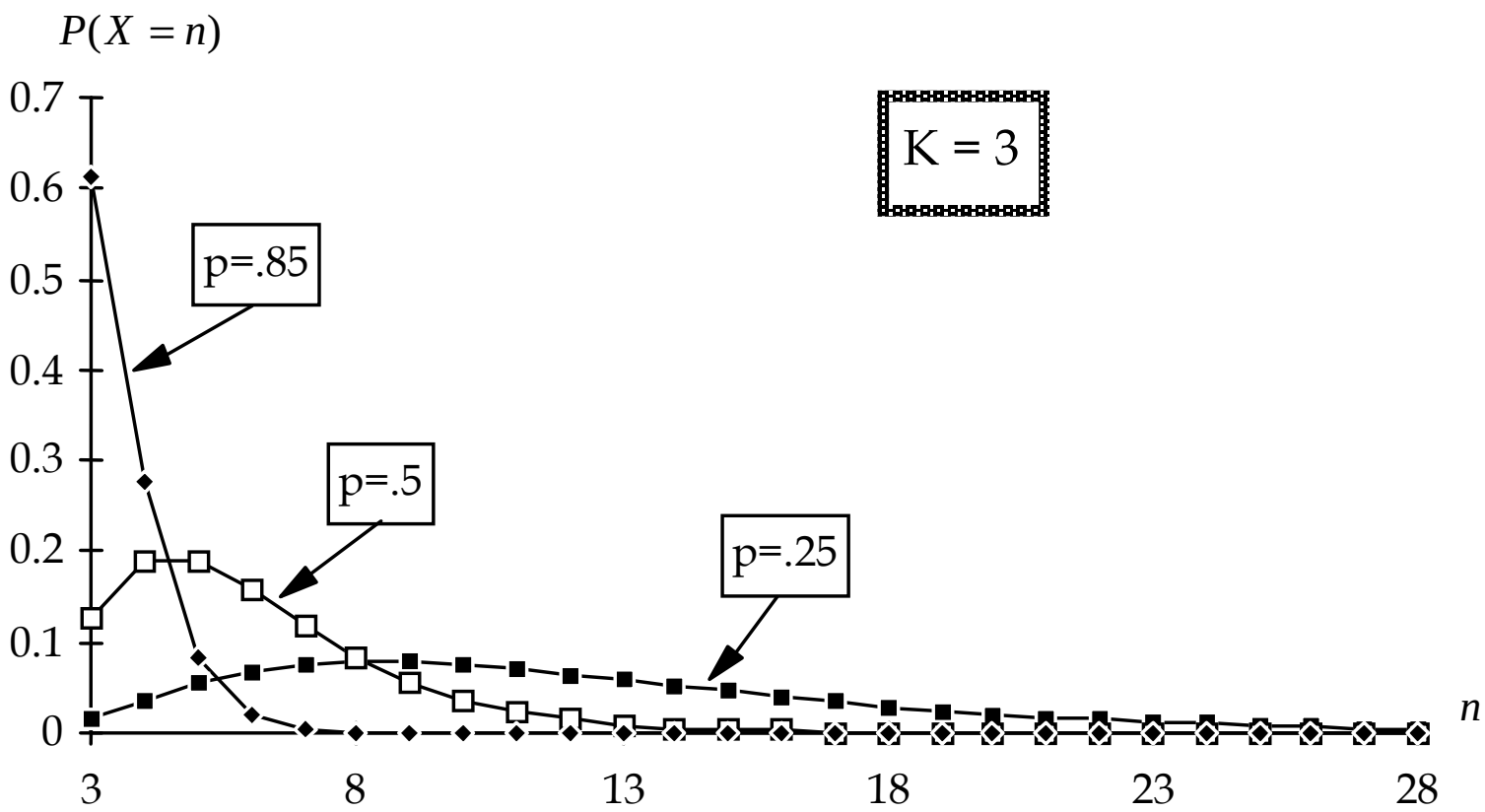
$$P(X = n) = \binom{n-1}{k-1} p^k (1-p)^{n-k}; \; k \leq n < \infty$$

The **mean and variance** of the NB distribution can be computed easily by noting that a negative binomial random variable with parameters *k* and *p* is just the sum of *k* independent geometric random variables with parameter *p*.

$$W_i \rightarrow G(p); 1 \leq i \leq k$$

$$X = \sum_{i=1}^{k} W_i \qquad E(X) = \sum_{i=1}^{k} E(W_i) = \frac{k}{p} \qquad D(X) = \sum_{i=1}^{k} D(W_i) = \frac{k(1-p)}{p^2}$$

Example: The following chart illustrates some examples of negative binomial distributions for k=3. Notice that for values of p near 0.5, the distribution has a single mode near 5. This mode moves away from the origin and diminishes in magnitude as p decreases indicating an increase in variance for small p. The negative binomial distribution has a shape similar to the geometric distribution for large values of p.



***Note*:**

Comparison of Binomial and Negative Binomial Distributions

<u>Binomial distribution</u>

$$P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}; 0 \le k \le n$$

For the binomial, the number of trials (n) is fixed and the number of successes (k) is random variable.

<u>Negative binomial distribution</u>

$$P(X = n) = \binom{n-1}{k-1} p^k (1-p)^{n-k}; k \le n < \infty$$

For the negative binomial, the number of successes (k) is fixed and the number of trials (n) is variable.

Poisson process

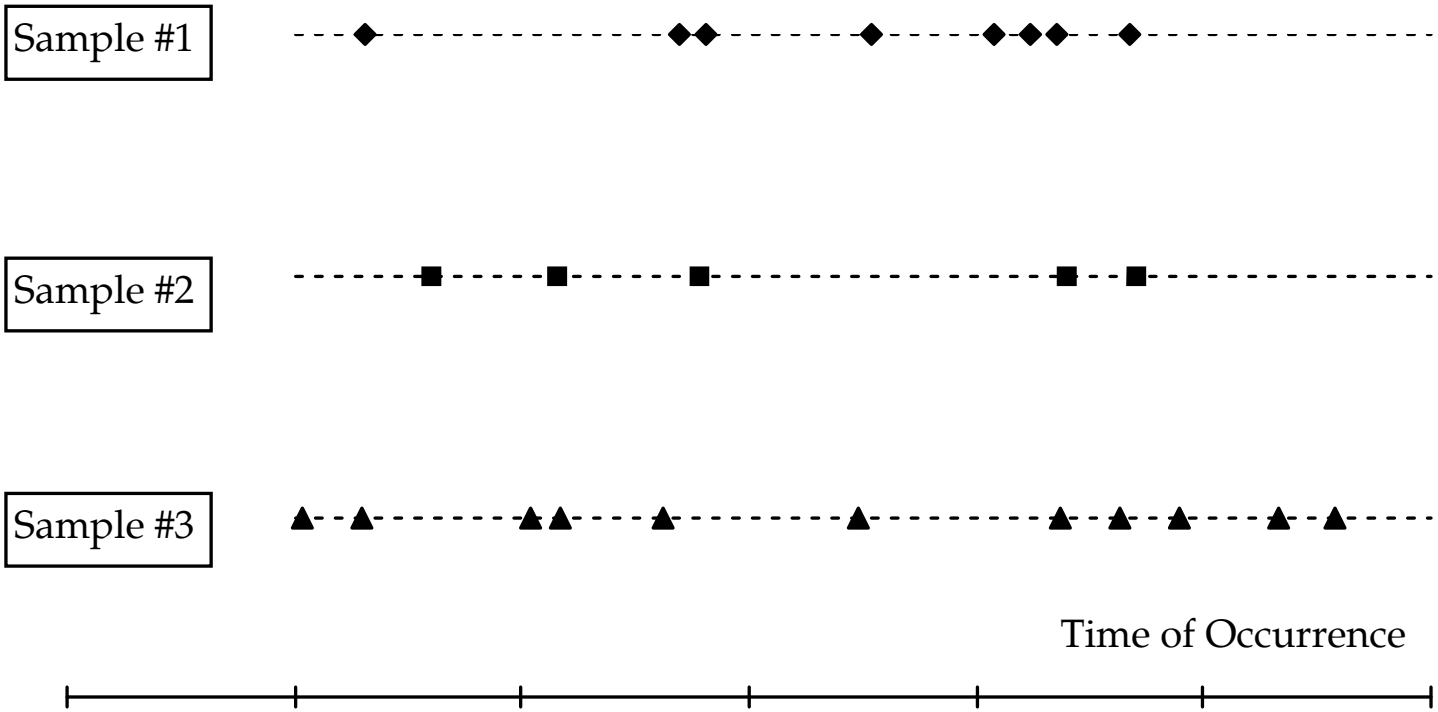
The Poisson process is a second general type of sample space model which is widely applied in practice. The Poisson process may be viewed as the continuous time generalization of a sequence of Bernoulli trials, sometimes called the Bernoulli process. The Poisson process describes the sample space of randomly occurring events in some time interval. The Poisson process assumes that the **rate at which events occur is constant** throughout the interval or region of observation and those events occur **independently** of each other.

λ … lambda $\rightarrow$ the rate at which events occur (proportional to probability of occurrence of one event per unit time)

However, events occurring over some region can also be modeled by a Poisson process.

Examples:

- Customers arriving at a bank to transact business.
- patients arriving at a clinic for treatment
- telephone inquiries received by a government office, *etc.*



Poisson distribution

The Poisson distribution has a single parameter and therefore we denote this random variable by the symbolic notation,

$$X \rightarrow P(\lambda t)$$

Consider a Poisson process that is observed for a time period t. Suppose the rate of occurrence of events is λ during the time period. Then the total rate of occurrence over the entire observation interval is λt. Now divide the interval t into n subintervals of equal length t/n. Occurrence of events in each of these intervals will be mutually independent at constant rate $\lambda t/n$. If n becomes large enough, the interval lengths, t/n, will become small enough that the probability of two events in one interval is effectively zero and the probability of one event is proportional to $\lambda t/n$. Then the distribution of the number of events in the total interval t can be approximated by the distribution of a binomial random variable with parameters n and $\lambda t/n$.

Thus,

$$P(X = k) = \binom{n}{k}\left(\frac{\lambda t}{n}\right)^k\left(1-\frac{\lambda t}{n}\right)^{n-k}$$

Taking the limit as *n* goes to infinity, this expression becomes

$$\lim_{n\to\infty}\binom{n}{k}\left(\frac{\lambda t}{n}\right)^k\left(1-\frac{\lambda t}{n}\right)^{n-k} = \frac{(\lambda t)^k}{k!}\lim_{n\to\infty}\frac{n!}{(n-k)!\,n^k}\left(1-\frac{\lambda t}{n}\right)^{n-k} = \frac{(\lambda t)^k e^{-\lambda t}}{k!} \quad ,$$

We can express the distribution of a Poisson random variable as:

$$P(X = k) = \frac{(\lambda t)^k e^{-\lambda t}}{k!};\ 0 \le k < \infty$$

Poisson variable is normed:

$$\sum_{k=0}^{\infty} P(X=k) = \sum_{k=0}^{\infty} \frac{(\lambda t)^k e^{-\lambda t}}{k!} = e^{-\lambda t} \sum_{k=0}^{\infty} \frac{(\lambda t)^k}{k!} = 1$$

Calculation of M**ean**:

$$E(X) = \sum_{k=0}^{\infty} kP(X=k) = \sum_{k=0}^{\infty} k \frac{(\lambda t)^k e^{-\lambda t}}{k!} = \lambda t \sum_{k=1}^{\infty} \frac{(\lambda t)^{k-1} e^{-\lambda t}}{(k-1)!} = \lambda t$$

Variance:

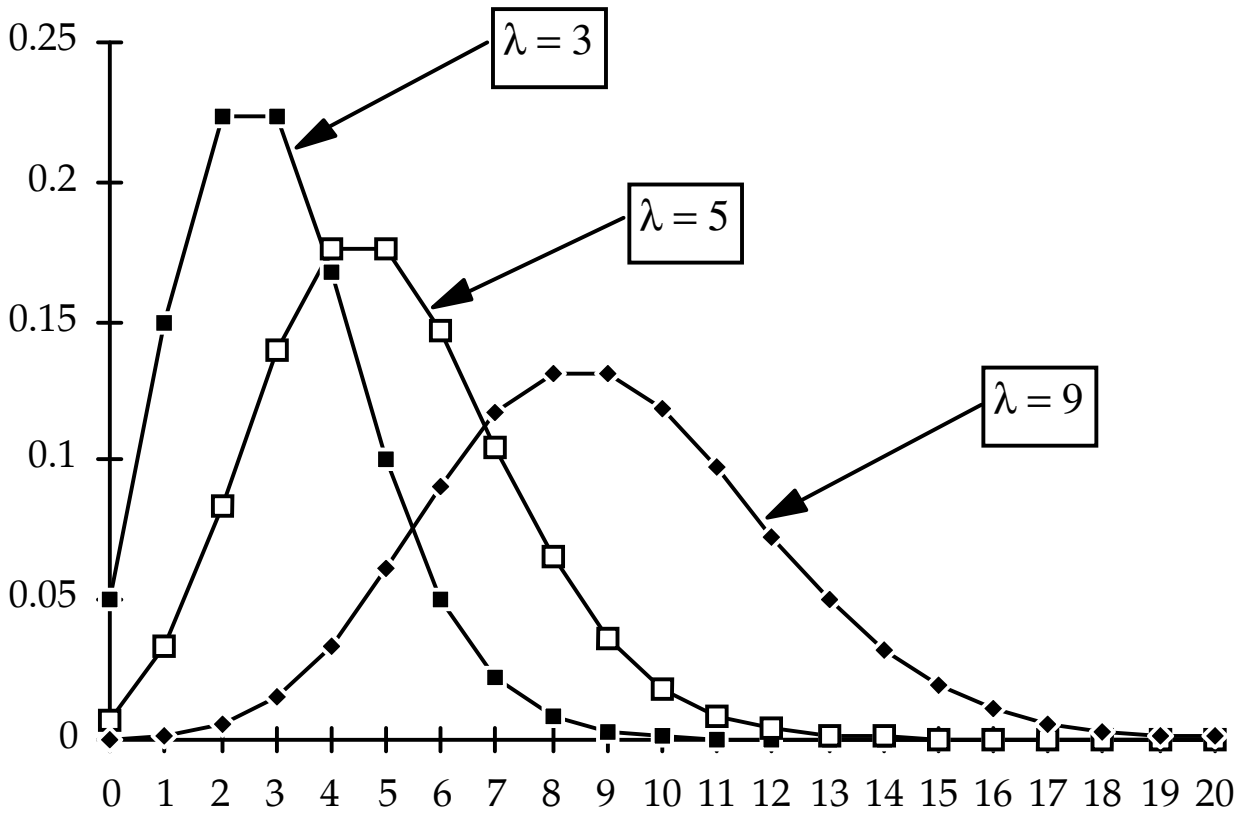
$$D(X) = E[(X - E(X)]^2) = \mathrm{E}(\mathrm{X}^2) - (\mathrm{E}(\mathrm{X}))^2 = \lambda t$$

$$E(X^2) = \sum_{k=0}^{\infty} k^2 P(X=k) = \sum_{k=2}^{\infty} k(k-1) \frac{(\lambda t)^k e^{-\lambda t}}{k!} + \mathrm{E}(\mathrm{X}) =$$

$$= (\lambda t)^2 \sum_{k=2}^{\infty} \frac{(\lambda t)^{k-2} e^{-\lambda t}}{(k-2)!} + \lambda t = (\lambda t)^2 + \lambda t$$

Example:

Some examples of Poisson distributions are illustrated below. Notice that at the value $\lambda = 9$, the distribution becomes almost symmetric.



Continuous Probability Distributions

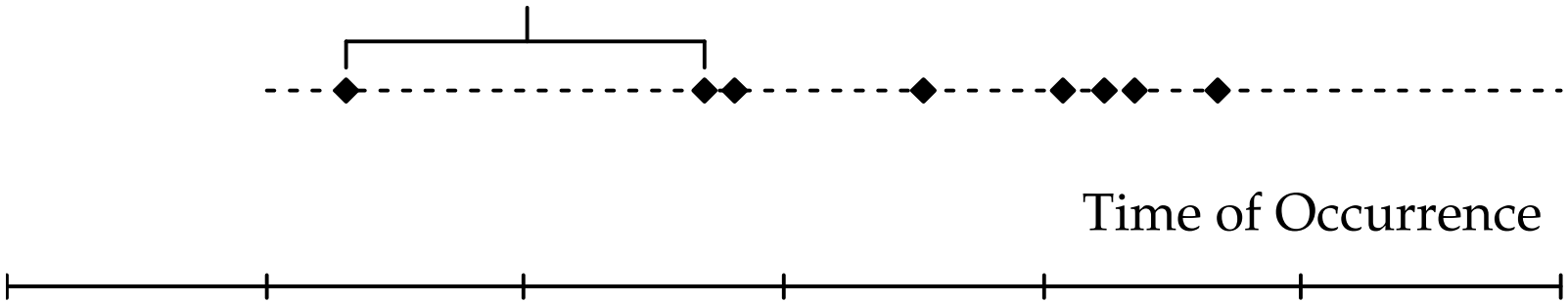
Aim - you should be able to

- characterize types of continuous distributions: exponential, Gamma and Weibull
- describe context between continuous distributions

Exponential Distribution

$$T \rightarrow E(\lambda)$$

T = Time Between Events



T has an Exponential Distribution

Derivation of *pdf*

N_t … number of Poisson events in interval $(0,t)$ $N_t \rightarrow P(\lambda t)$

Consider the equivalency of two events: $N_t \geq 1 \Leftrightarrow T < t$

$$F(t)=P(T < t) = P(N_t \geq 1) = 1 - P(N_t < 1)$$
$$= 1 - e^{-\lambda t}$$

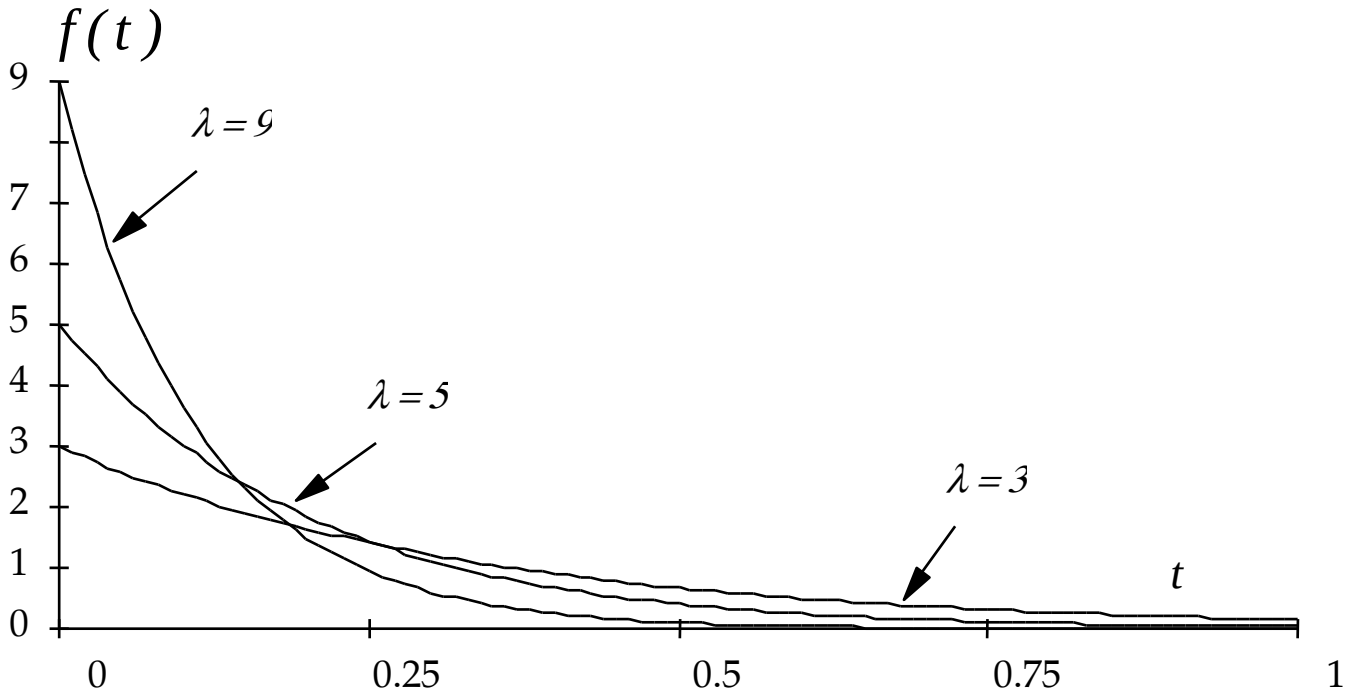
pdf: $$f(t) = \lambda e^{-\lambda t}; t \geq 0$$

$$E(T) = \int_{t=0}^{\infty} \lambda t\, e^{-\lambda t}\, dt = \frac{1}{\lambda}$$

$$DT = ET^2 - (ET)^2 = \ldots = \frac{1}{\lambda^2}$$

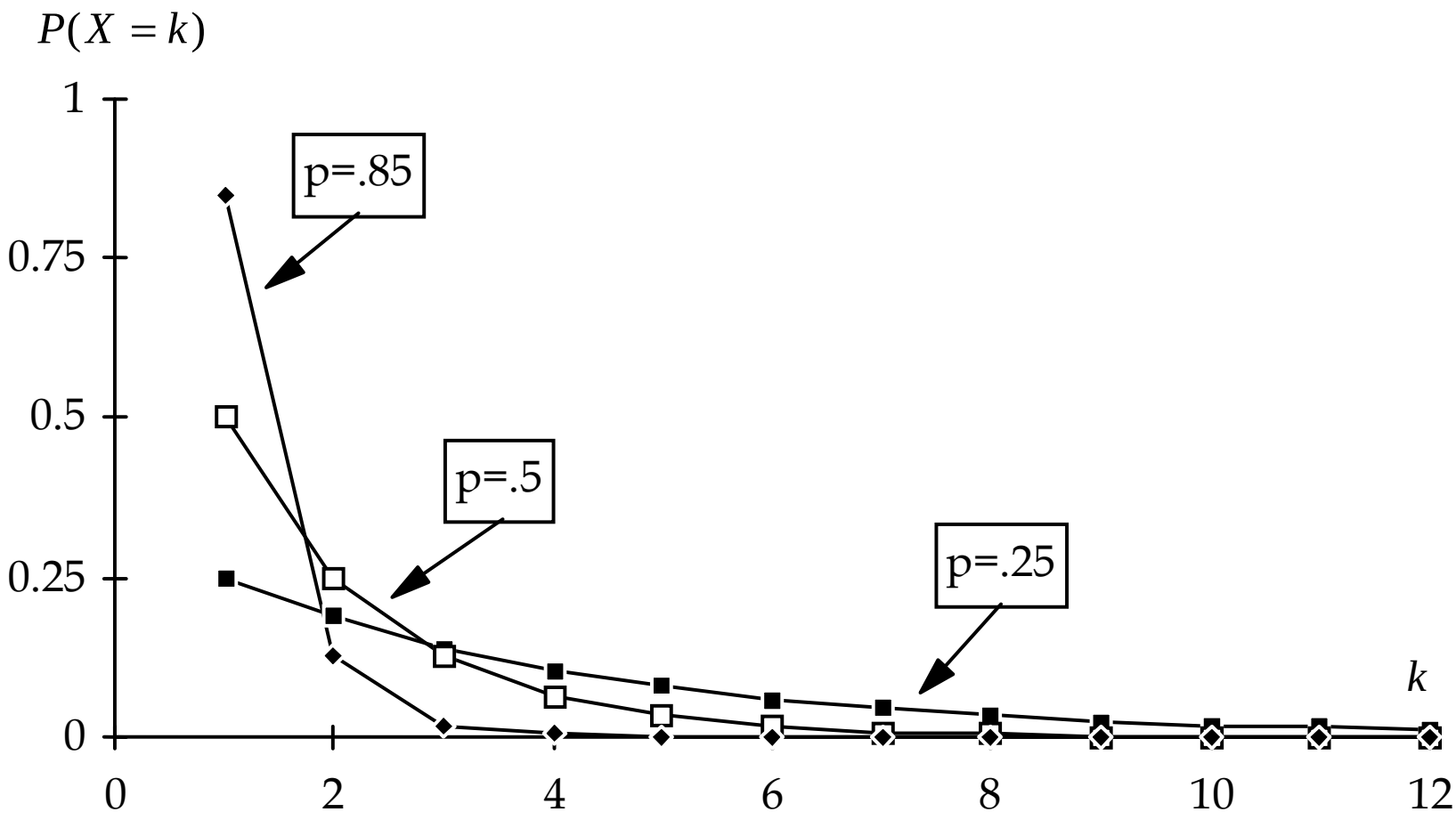
Example:

The following graph illustrates some examples of the probability density functions of exponential random variables. Notice that the shape of the exponential density is similar to the shape of the geometric probability distribution.



probability density function of exponential distribution

Compare with: **Geometric distribution:** $X \rightarrow G(p)$



Hazard function

X … non-negative RV, continuous

Definition: we call the function $\lambda(t)$ **as hazard function**

$$\lambda(t) = \frac{f(t)}{1 - F(t)} \qquad F(t) < 1.$$

Probability meaning of hazard function:

$P(t < X \le t + \Delta t | X > t) \approx \dfrac{f(t)}{1 - F(t)} \Delta t = \lambda(t).\Delta t$, see below.

If X … time to failure of an equipment, then hazard function represents conditional probability (is proportional), that if till time t no failure occurred, it occurs in following time interval Δt:

$$P\left(t < X < t+\Delta t \middle| X \geq t\right) = \frac{P\left(t < X < t+\Delta t, X \geq t\right)}{P\left(X \geq t\right)} = \frac{P\left(t < X < t+\Delta t\right)}{P\left(X \geq t\right)}$$

$$= \frac{1}{1-F\left(t\right)} \cdot \frac{F\left(t+\Delta t\right)-F\left(t\right)}{\Delta t} \cdot \Delta t$$

For $\Delta t \to 0$ we obtain $\dfrac{F\left(t+\Delta t\right)-F\left(t\right)}{\Delta t} \to \dfrac{dF}{dt} = f\left(t\right)$,

So that:

$$P\left(t < X < t+\Delta t \middle| X \geq t\right) \approx \frac{f\left(t\right)}{1-F\left(t\right)} \Delta t = \lambda\left(t\right).\Delta t$$

Hazard function is local reliability characteristics. It represents approximately the probability that a object which is right functioning till t, will fail in $(t, t + 1)$

Hazard function 2

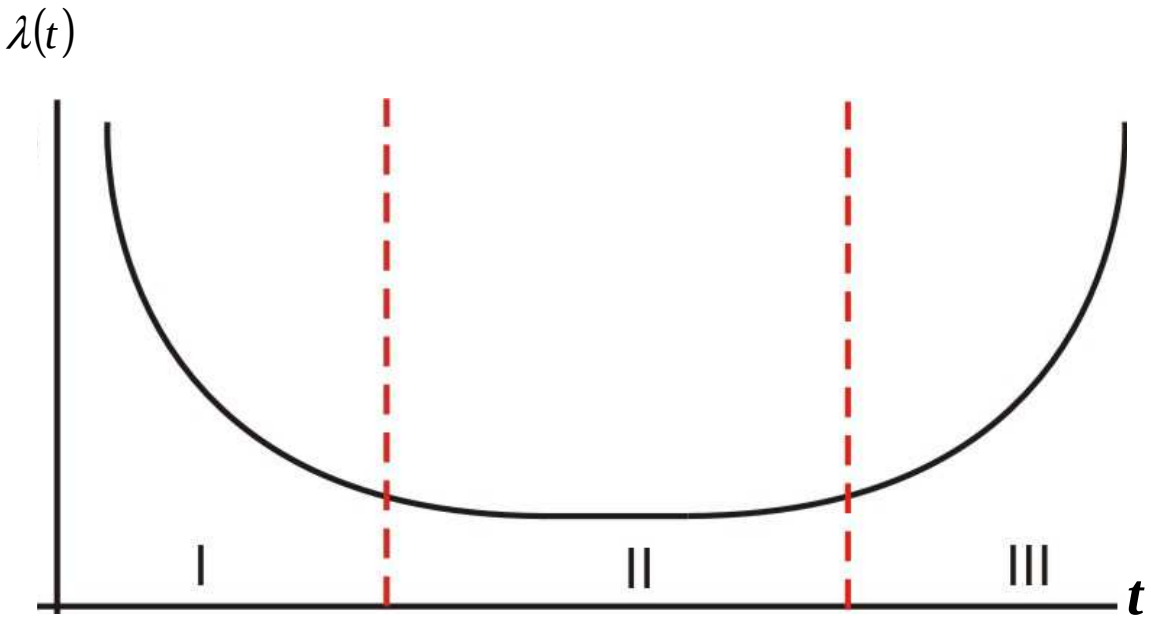
Relations between $f(t), F(t), \lambda(t)$:

	$F(t)$	$f(t)$	$\lambda(t)$
$F(t)$	$F(t)$	$\int_0^t f(x)dx$	$1-exp[-\int_0^t \lambda(x)dx]$
$f(t)$	$\frac{dF(t)}{dt}$	$f(t)$	$\lambda(t)\cdot exp[-\int_0^t \lambda(x)dx]$
$\lambda(t)$	$\frac{\frac{dF(t)}{dt}}{1-F(t)}$	$\frac{f(t)}{1-\int_0^t f(x)dx}$	$\lambda(t)$

Hazard function 3

Graphics demonstration of hazard function:

Bathtub curve (I, II, III).



I …period of early failures

II … steady state period

III …period of ageing or wear-out period

Hazard function 4

$$h(t) = \frac{f(t)}{1 - F(t)} \quad \text{, if} \quad F(t) < 1$$

Exponential distribution:

$$h(t) = \frac{f(t)}{1 - F(t)} = \frac{\lambda e^{-\lambda t}}{1 - (1 - e^{-\lambda t})} = \lambda = konst.$$

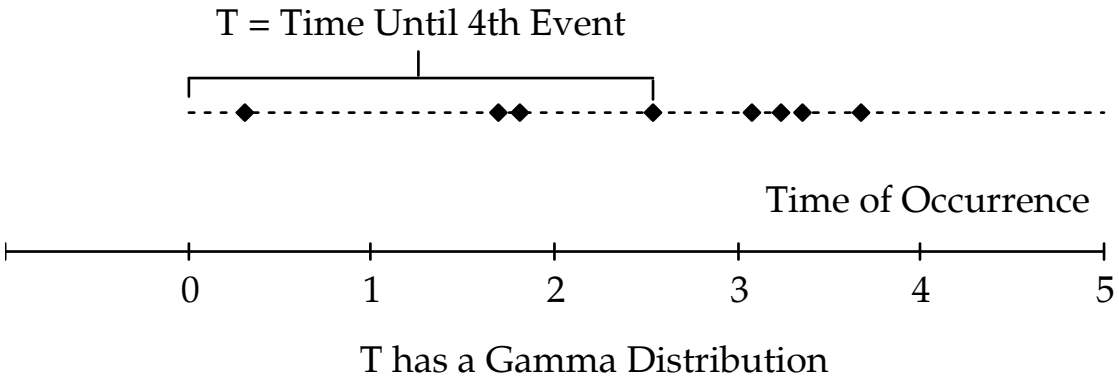
$\Rightarrow$ the "**no memory**" property of the exponential distribution

Gamma distribution 1

$$T \rightarrow Ga\,(k, \lambda)$$

T ... time of occuring of kth event in Poisson process

Example: $k = 4$:



Gamma distribution 2

The gamma distribution function for any integer value of k can be derived by the following way. Since the gamma arises as the sum of k independent, identically distributed exponential random variables, the distribution function of the gamma is the probability that the **sum of *k* exponentials** is less than or equal to some value t. This implies that there have been at least k occurrences of a Poisson process within time t, the probability of which is given by the cumulative distribution of a Poisson random variable with rate parameter λt, where λ is the rate of the underlying Poisson process.

Gamma distribution:

$\{N_t \geq k\} \Leftrightarrow \{(T_k = \sum_{i=1}^{k} X_i) < t\}$ equivalency of two events, $N_t \rightarrow P(\lambda t)$

$T_k = \sum_{i=1}^{k} X_i$ .... Gama RV = sum of k exponentially distributed times:

$T_k = X_1 + X_2 + X_3 + \ldots + X_k$

Distribution function: $F(t) = P(T_k < t) =$

$$P\left(\sum_{i=1}^{k} X_i < t\right) = P(N_t \geq k) = 1 - P(N_t < k) = 1 - \sum_{j=0}^{k-1} e^{-\lambda t} \cdot \frac{(\lambda t)^j}{j!} = 1 - e^{-\lambda t}\left[\sum_{j=0}^{k-1} \frac{(\lambda t)^j}{j!}\right]$$

Gamma distribution 3

Derivation of *F(t)* gives *pdf*:

$$f(t) = \lambda e^{-\lambda t}\left[\sum_{j=0}^{k-1}\frac{(\lambda t)^j}{j!}\right] - e^{-\lambda t}\left[\sum_{j=1}^{k-1}\frac{\lambda(\lambda t)^{j-1}}{(j-1)!}\right]$$

$$= \lambda^k e^{-\lambda t}\left[\frac{t^{k-1}}{(k-1)!}\right]; \quad t > 0$$

$$T_k = X_1 + X_2 + X_3 + \ldots + X_k$$

$$X_i \rightarrow E(\lambda)$$

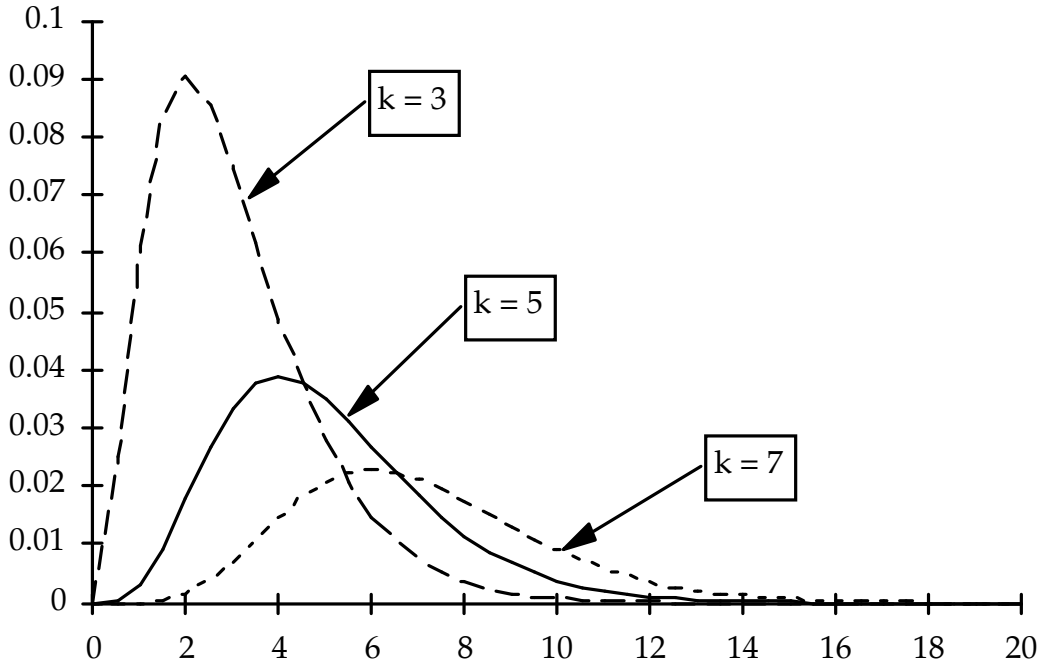
$$ET_k = EX_1 + EX_2 + \ldots + EX_k = \frac{1}{\lambda} + \ldots + \frac{1}{\lambda} = \frac{k}{\lambda} \quad \text{... mean}$$

$$DT_k = \ldots. = \frac{k}{\lambda^2} \quad \text{.... variance of } T_k$$

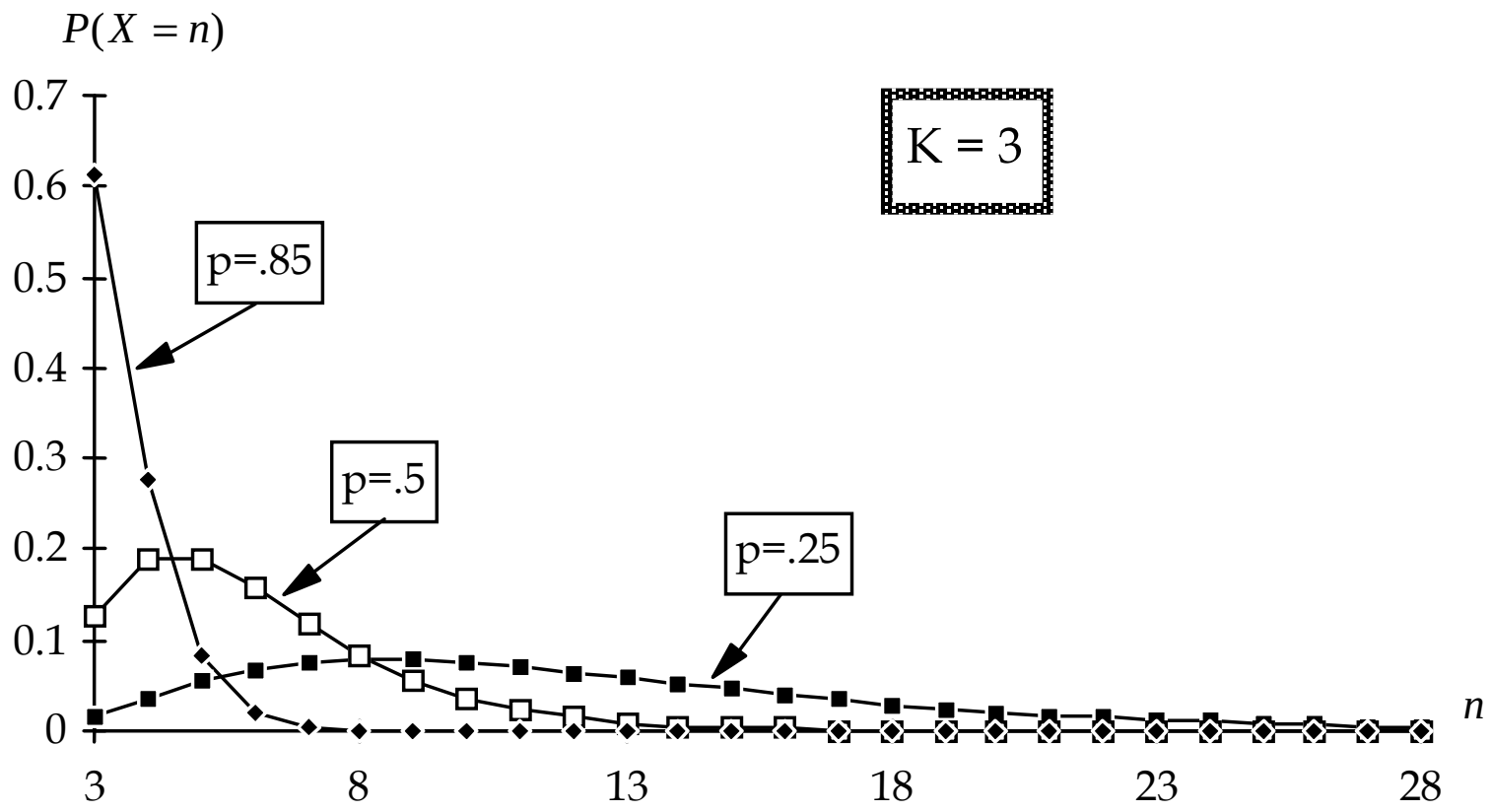
If *k* is integer, we speak about Erlang distribution.

Example:

Examples of gamma probability density functions for $\lambda = 1$. Notice that the gamma density has a single mode which moves away from the origin as k increases. Also the dispersion increases and the distribution become more symmetric and k increases.

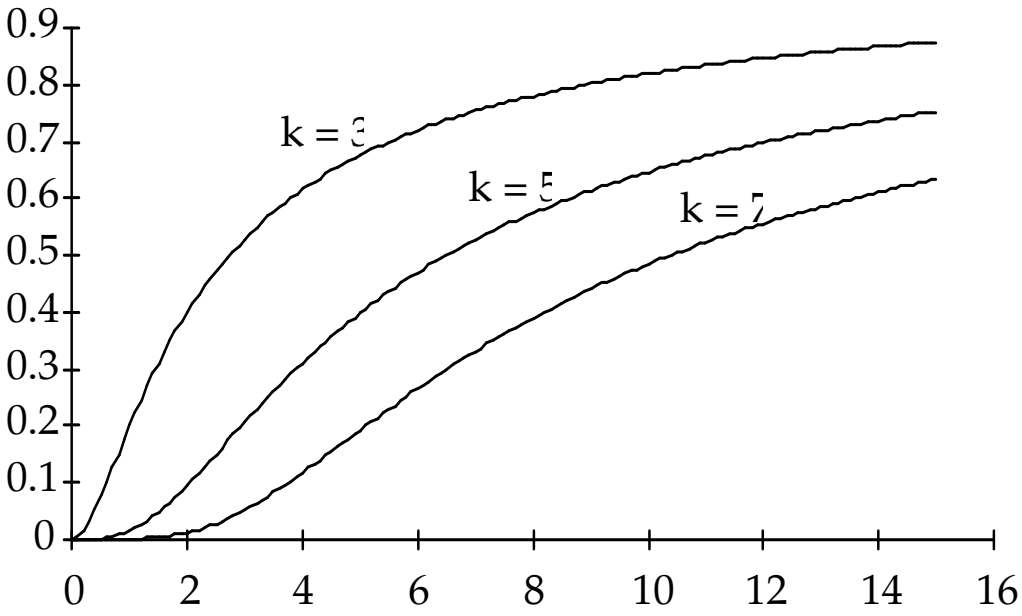


Comparison with **Negative binomial random variable** $X \rightarrow NB(k, p)$



Hazard function:

$$h(x) = \frac{\lambda}{(k-1)!\sum_{j=0}^{k-1} \frac{1}{(k-1-j)!(\lambda x)^j}}$$



Hazard function, for λ=1

Weibull distribution

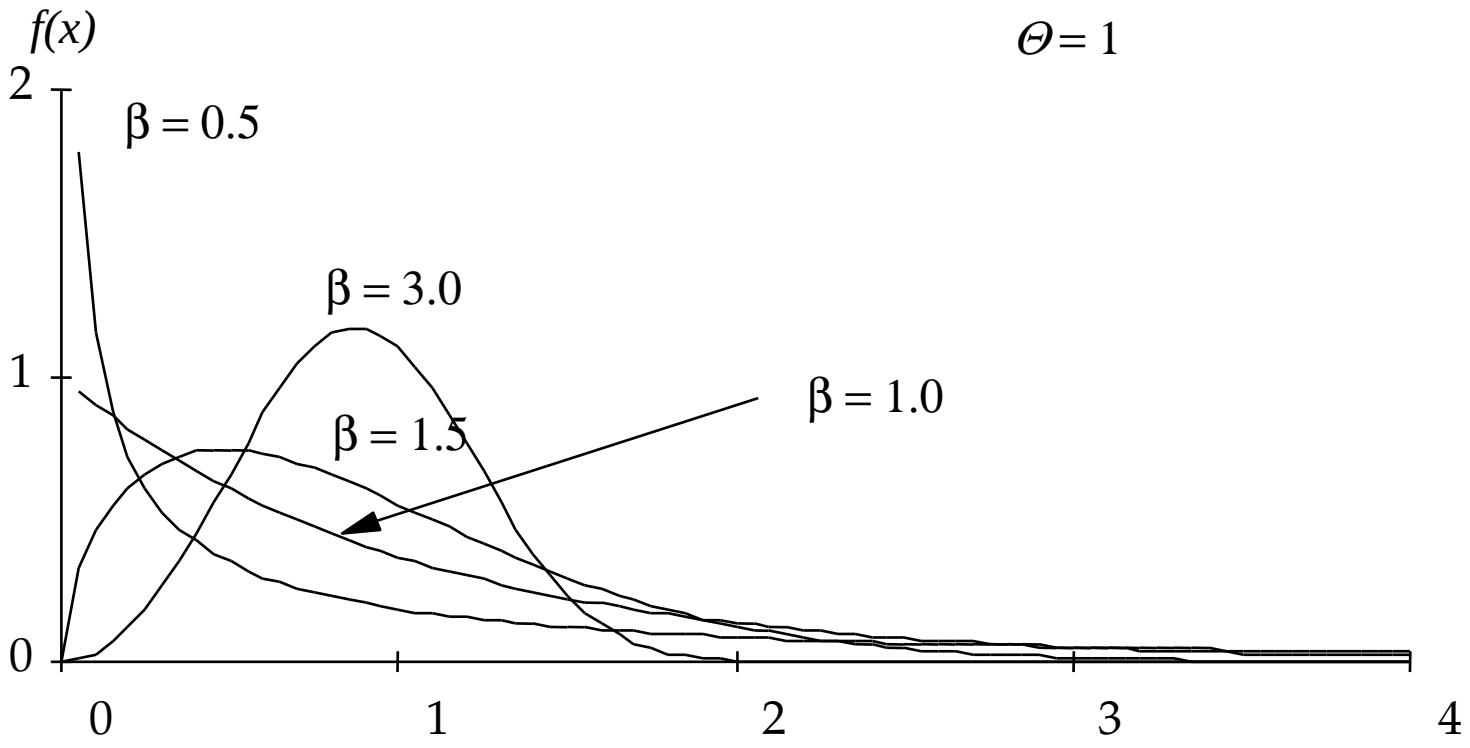
Distribution function:

$$F(x) = 1 - e^{-\left(\frac{x}{\Theta}\right)^{\beta}} \quad , \quad \Theta > 0, \beta > 0, x > 0$$

β ... shape parameter, Θ ... scale parameter .

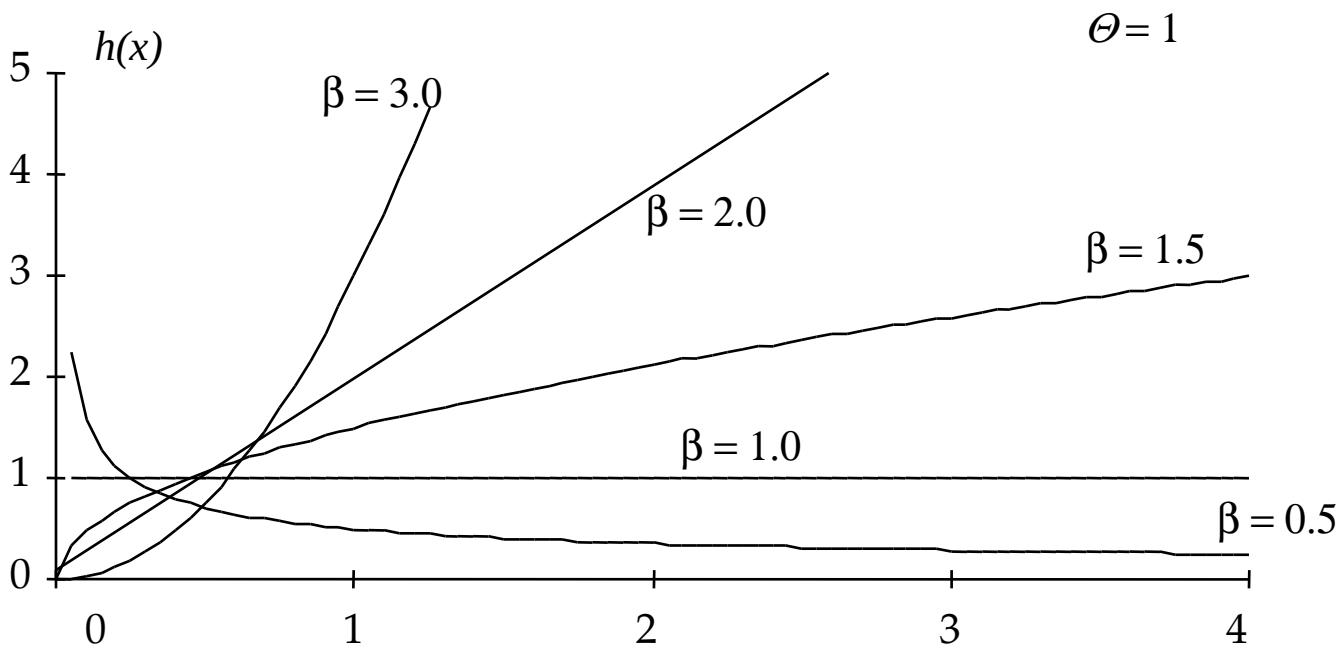
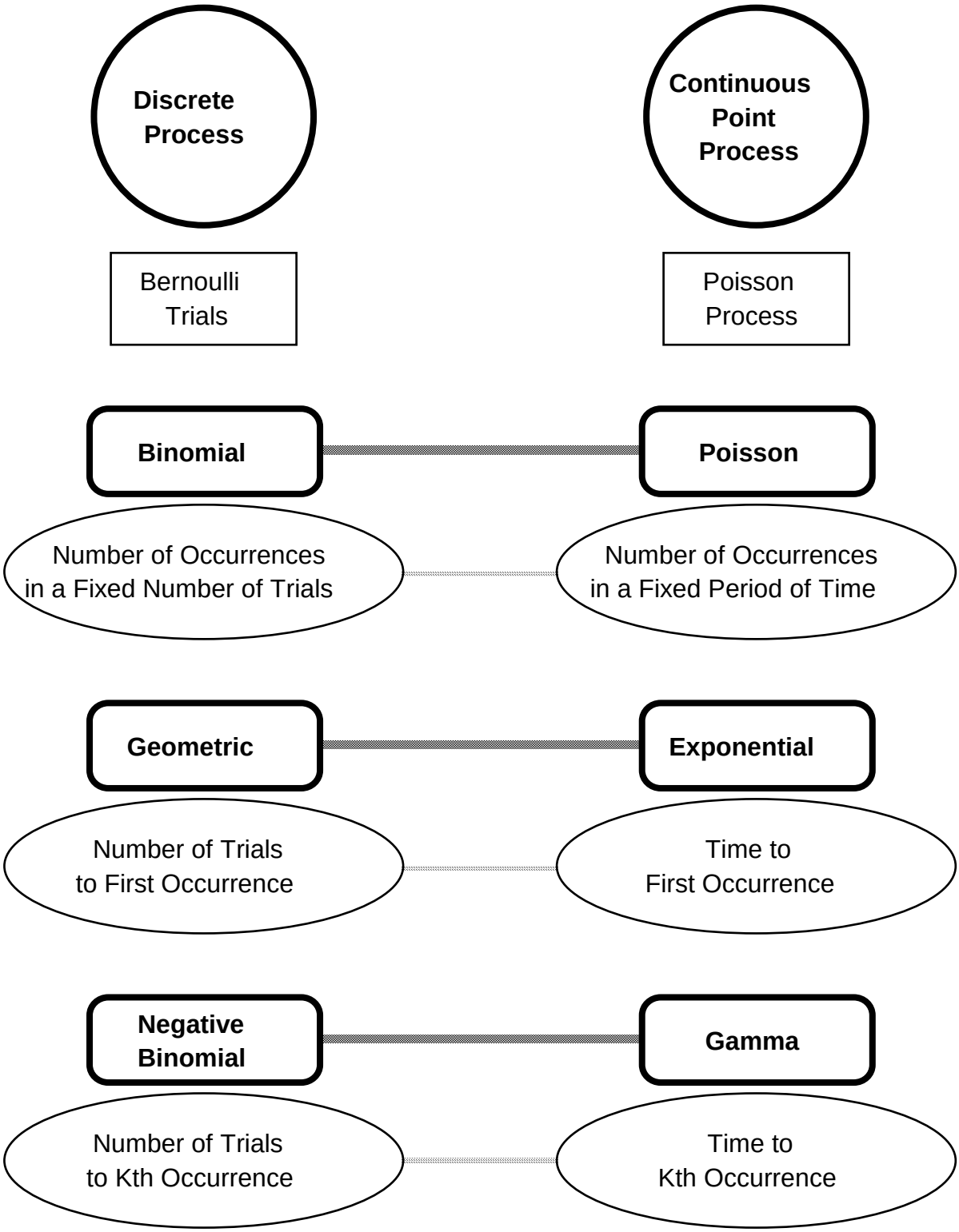
pdf: $$f(x) = \frac{\beta}{\Theta}\left(\frac{x}{\Theta}\right)^{\beta-1} e^{-\left(\frac{x}{\Theta}\right)^{\beta}}$$

Some examples of the Weibull density:



Hazard function: $$h(x) = \frac{\beta}{\Theta} \cdot \left(\frac{x}{\Theta} \right)^{\beta - 1}$$

The Weibull distribution is very flexible and we use it in Reliability theory for modeling of the random variable "**time to failure**".



Relations between distributions