BITS Pilani
K K Birla Goa Campus

Probability & Statistics,

Dr. Jajati Keshari Sahoo
Department of Mathematics

Gamma Distribution

The probability density of the **gamma distribution** is given by

$$f(x) = \begin{cases} \dfrac{1}{\Gamma(\alpha)\beta^{\alpha}} x^{\alpha-1} e^{-x/\beta} & \text{for } x > 0,\ \alpha > 0,\ \beta > 0 \\ 0 & \text{elsewhere} \end{cases}$$

where $\Gamma(\alpha)$ is a value of the **gamma function**, defined by

$$\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} e^{-x} dx$$

Gamma Distribution

Properties:

- $\Gamma(\alpha) = (\alpha - 1)\Gamma(\alpha - 1)$ for any $\alpha > 1$.

  Proof:
  $$\Gamma(\alpha) = -e^{-x}x^{\alpha-1}\Big|_0^\infty - \int_0^\infty [-e^{-x}(\alpha-1)x^{\alpha-2}$$
  $$= 0 + (\alpha-1)\int_0^\infty e^{-x}x^{\alpha-2} = (\alpha-1)\Gamma(\alpha-1)$$

- $\Gamma(1) = 1$ and $\Gamma(1/2) = \sqrt{\pi}$

- $\Gamma(\alpha) = (\alpha - 1)!$ if α is a positive integer

Gamma Distribution

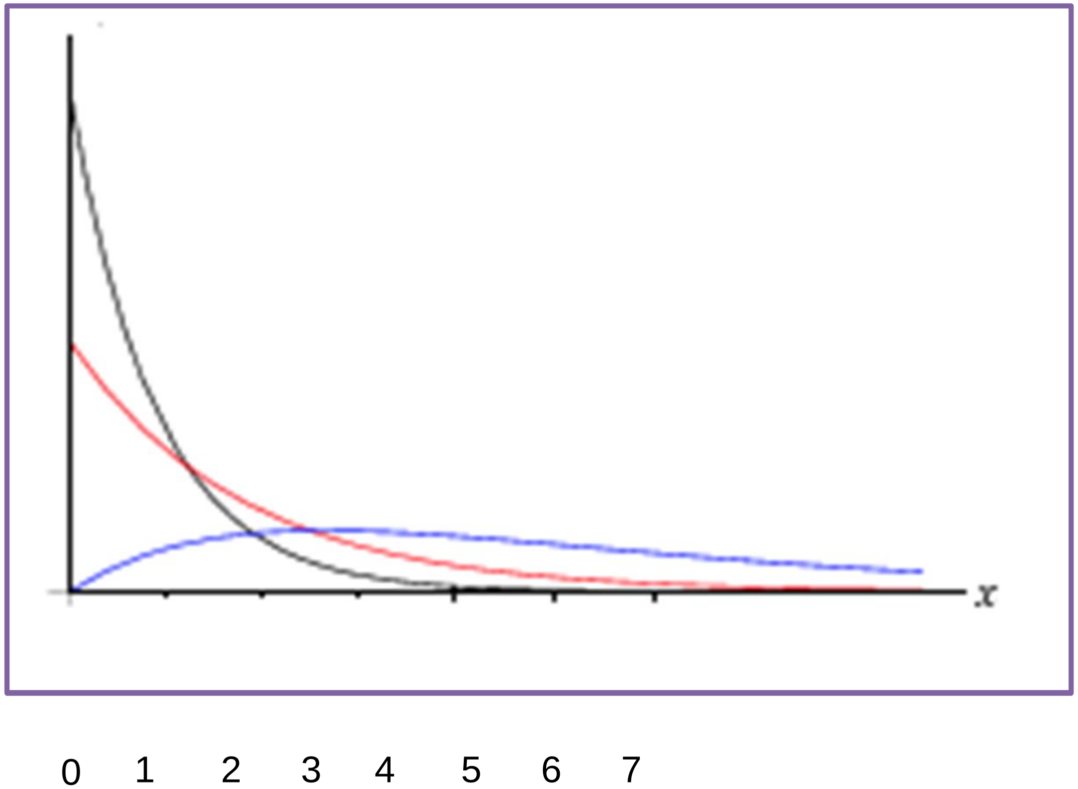


Figure: Graph of some gamma probability density functions

Gamma Distribution

Mean of gamma distribution:

$$\mu = \alpha\beta$$

Proof:

$$\mu = \frac{1}{\beta^{\alpha}\Gamma(\alpha)}\int_0^{\infty} x.x^{\alpha-1}e^{-x/\beta}dx$$

(put $y = x/\beta$)

$$= \frac{\beta}{\Gamma(\alpha)}\int_0^{\infty} y^{\alpha}e^{-y}dy = \frac{\beta\,\Gamma(\alpha+1)}{\Gamma(\alpha)}$$

$$= \alpha\beta$$

Gamma Distribution

Variance of gamma distribution:

$$\sigma^2 = \alpha\beta^2$$

Proof:

$$\mu_2 = \frac{1}{\beta^\alpha \Gamma(\alpha)} \int_0^\infty x^2 . x^{\alpha-1} e^{-x/\beta} dx$$

$$= \frac{\beta^2}{\Gamma(\alpha)} \int_0^\infty y^{\alpha+1} e^{-y} dy = \frac{\beta^2 \, \Gamma(\alpha+2)}{\Gamma(\alpha)}$$

$$= \alpha\beta^2(\alpha+1)$$

Hence $\sigma^2 = \mu_2 - \mu^2 = \alpha\beta^2(\alpha+1) - \alpha^2\beta^2 = \alpha\beta^2.$

Gamma Distribution

MGF of Gamma Distribution:

$$M_X(t) = E(e^{Xt}) = \frac{1}{\beta^{\alpha}\Gamma(\alpha)}\int_0^{\infty} e^{xt}x^{\alpha-1}e^{-x/\beta}dx$$

$$= \frac{1}{\beta^{\alpha}\Gamma(\alpha)}\int_0^{\infty} x^{\alpha-1}e^{-x(1/\beta - t)}dx$$

$$= \frac{1}{\beta^{\alpha}\Gamma(\alpha)}\frac{\Gamma(\alpha)}{(1/\beta - t)^{\alpha}}, \quad t < 1/\beta$$

$$= \frac{1}{(1-t\beta)^{\alpha}}, \quad t < 1/\beta$$

Gamma Distribution

Problem 1

If $X_1, X_2, \cdots, X_n$ are independent random variables and follows gamma distribution with parameters $(\alpha_1, \beta), (\alpha_2, \beta), \cdots, (\alpha_n, \beta)$ respectively. Show that $X_1 + X_2 + \cdots + X_n$ follows gamma distribution with parameter $(\alpha_1 + \alpha_2 + \cdots + \alpha_n, \beta)$.

Exponential Distribution

The density function of exponential distribution is given by

$$f(x) = \begin{cases} \dfrac{1}{\beta} e^{-x/\beta} & \text{for } x > 0,\ \beta > 0 \\ 0 & \text{elsewhere} \end{cases}$$

Note: Exponential is the special case of gamma distribution where $\alpha = 1$.

Mean, variance and generating function of the exponential distribution:

$$\mu = \beta,\ \sigma^2 = \beta^2 \text{ and } M_X(t) = \frac{1}{1-\beta t}$$

Exponential Distribution

Problem 2

If X is exponential random variable with parameter β. Find the cumulative distribution function of X.

Ans: $F(x) = 1 - e^{-x/\beta}$

Exponential Distribution

Problem 3

If $X_1, X_2, \cdots, X_n$ are independent exponential distributions with parameters $\beta_1, \beta_2, \cdots, \beta_n$ respectively. Find the density function of the random variable $Y = \min(X_1, X_2, \cdots, X_n)$.

Memory Loss Property

Problem 4

If X is exponential random variable with parameter β. Show that

$P(X > x+t \mid X > t) = P(X > x)$ for $t > 0$.

Note: The converse also true. Verify ?

Memory Loss Property

Remark:

- If X represents the lifetime of a device, then memory loss property states that if the device has been working for time t, then the probability that it will survive an additional time x depends only on x (not on t) and is identical to the probability of survival for time x of a new device.

- The equipment does not remember that it has been in use for time t.

Memory Loss Property

Problem 5

Suppose the life length of a machine has an exponential distribution with $\beta = 10$ years. A 7 years used machine is bought by someone. Find the probability that it will not fail in the next 5 years.

Exponential Distribution

Problem 06

Personnel of a company use an online terminal to make routine calculations. If the time each person spends in a session at a terminal has an exponential distribution with an **average value of 36 minutes,** find the probability that a person

(i) Will spend 30 minutes or less at the terminal

(ii) If he has already been at the terminal for 30 minutes, what is the probability that he will spend more than another hour at the terminal.

Application Exponential Distribution

Theorem:

If in a Poisson process average number of arrivals per unit time is λ. Let W denote waiting time between successive arrivals (or the time until the first arrival). Then W has an exponential distribution with

$$\beta = \frac{1}{\lambda}.$$

Application Exponential Distribution

Proof:

Let X be the number of arrivals in the time interval of length t.

$\Rightarrow$ X follows Poisson process and

$$P(X = x) = \frac{(\lambda t)^x e^{-\lambda t}}{x!}, x = 0, 1, \cdots$$

Application Exponential Distribution

Proof(cont.):

$$P(\text{waiting time between successive arrivals be at least } t)$$

$$= P(W > t)$$

$$= P(\text{no arrivals during a time interval of length } t)$$

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

$$\Rightarrow P(\text{waiting time between successive arrivals } < t)$$

$$= 1 - e^{-\lambda t} = F(t), \text{ the distribution function of W.}$$

Application Exponential Distribution

Proof(cont.):

So if waiting time between successive arrivals be random variable with the distribution function

$$F(t) = 1 - e^{-\lambda t}$$

So, the probability density of the waiting time between successive arrivals given by

$$f(t) = \frac{d}{dt}F(t) = \lambda e^{-\lambda t},$$

which is an exponential distribution with $\beta = \frac{1}{\lambda}.$

Problem 07

Given that the switchboard of consultant's office receives on the average 0.6 calls per minute, find the probabilities that the time between successive calls arriving at the switchboard of the consulting firm will be

(a) less than 1/2 minute;

(b) more than 3 minute.

Solution to Problem 07

Since λ = 0.6, the waiting time t between successive calls arriving at the switchboard, has an exponential distribution with β = 1/0.6, hence density function is given by

$$f(t) = \begin{cases} 0.6e^{-0.6t} & \text{for } t > 0 \\ 0 & \text{elsewhere} \end{cases}$$

$$\text{(a)}\ P(t < \frac{1}{2}) = \int_0^{1/2} 0.6e^{-0.6t}\,dt = -e^{-0.6t}\Big|_0^{1/2} = 1 - e^{-0.3}.$$

$$\text{(b)}\ P(t > 3) = \int_3^{\infty} 0.6e^{-0.6t}\,dt = -e^{-0.6t}\Big|_3^{\infty} = e^{-1.8}.$$

Problem 08

Customers arrive in a certain shop according to an approximate Poisson process at a mean rate of 20 per hour. What is the probability that the shopkeeper will have to wait more than 5 minutes for the arrival of the first customer? .

Problem 09

An average of 30 customers per hour arrive at a shop in accordance with a Poisson process. Find the probability that the shopkeeper will wait more than 5 minutes before the second customer arrives.

Chi-squared Distribution

This distribution is a special case of Gamma distribution in which $\alpha = \gamma/2$, where γ is a positive integer and $\beta = 2$.

Chi-squared Distribution

The density function of Chi-squared distribution is given by

$$f(x) = \begin{cases} \dfrac{1}{2^{\gamma/2}\Gamma(\gamma/2)} x^{\gamma/2-1} e^{-x/2}, & x > 0 \\ 0, & \textit{elsewhere} \end{cases}$$

We denote the random variable by X_{γ}^{2}

and γ is called degree of freedom.

Chi-squared Distribution

Mean, variance and moment generating function of the Chi-squared distribution:

$$E(X_{\gamma}^{2}) = \gamma$$

$$Var(X_{\gamma}^{2}) = 2\gamma$$

$$M_{X_{\gamma}^{2}}(t) = (1-2t)^{-\gamma/2}, \ t < 1/2$$

Advantage of Chi-squared distribution

Remark:

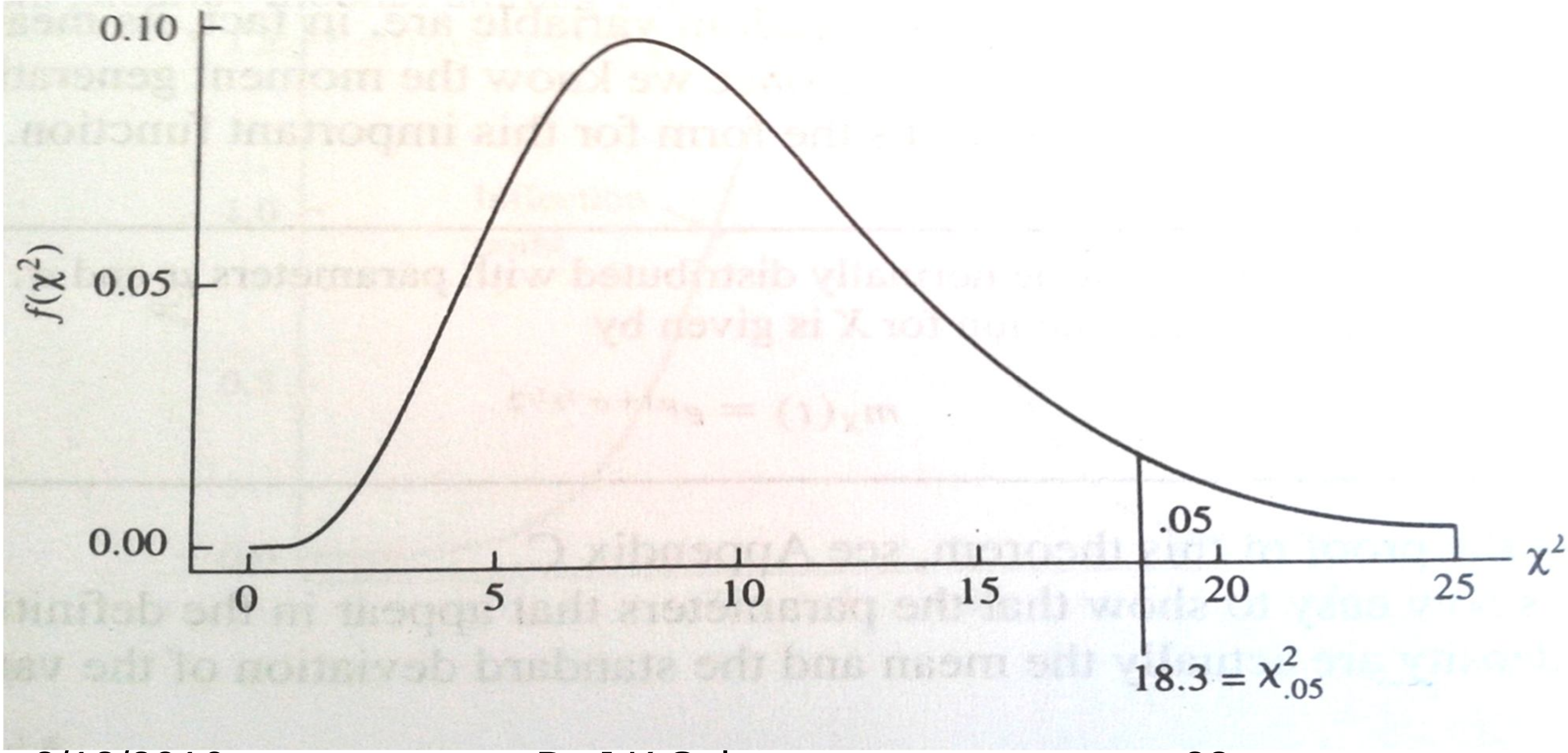
- Cumulative distribution values available for some points with specified degree of freedom.

- Sum of independent Chi-squared distributions is again Chi-squared.

Chi-squared Distribution

Notation:

$$P(X_{\gamma}^{2} \geq \chi^{2}_{\alpha,\gamma}) = \alpha$$



Chi-squared Distribution

How to use cumulative distribution Table:

$P[\chi^2_\gamma \leq t]$

γ \ F	0.005	0.010	0.025	0.050	0.100	0.250
1	0.0000393	0.000157	0.000982	0.00393	0.0158	0.102
2	0.0100	0.0201	0.0506	0.103	0.211	0.575
3	0.0717	0.115	0.216	0.352	0.584	1.21
4	0.207	0.297	0.484	0.711	1.06	1.92
5	0.412	0.554	0.831	1.15	1.61	2.67
6	0.676	0.872	1.24	1.64	2.20	3.45
7	0.989	1.24	1.69	2.17	2.83	4.25
8	1.34	1.65	2.18	2.73	3.49	5.07
9	1.73	2.09	2.70	3.33	4.17	5.90
10	2.16	2.56	3.25	3.94	4.87	6.74
11	2.60	3.05	3.82	4.57	5.58	7.58
12	3.07	3.57	4.40	5.23	6.30	8.44

Applications

- The aggregate insurance claims and the amount of rainfall accumulated in a reservoir are modeled by a gamma distribution.

- In neuroscience, the gamma distribution is often used to describe the distribution of inter-spike intervals.

- In genomics, the gamma distribution was applied in peak calling step (i.e. in recognition of signal) in ChIP-chip and ChIP-seq data analysis

Applications

- In bacterial gene expression, the copy number of a constitutively expresses protein often follows the gamma distribution, where the scale and shape parameter are, respectively, the mean number of bursts per cell cycle and the mean number of protein molecules produced by a single mRNA during its lifetime.

- It also used to fit life model of many devices based on certain assumptions.