San José State University

Math 161A: Applied Probability & Statistics

Special continuous distributions

Prof. Guangliang Chen

Section 4.3 The Normal distribution

Section 4.4 The Exponential and Gamma distributions

Introduction

In this lecture we cover the following special continuous distributions

- **Uniform**
- **Normal**
- **Exponential**
- **Gamma** (and Chi-square)

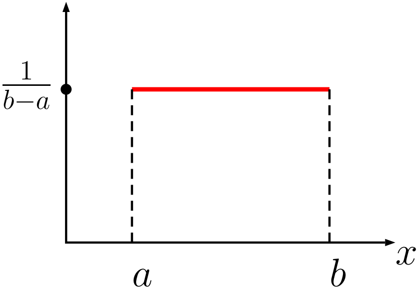
... by following the same treatment plan (as for discrete distributions):

- Examples
- Definition (via pdf)
- Expected value
- Variance
- Other useful properties (if any)

The Uniform distribution

Def 0.1 $(X \sim \text{Unif}(a, b))$**.** A continuous random variable X is said to have a **uniform** distribution with parameters a, b if it has the following probability density function (pdf):

$$f(x) = \begin{cases} \frac{1}{b-a}, & a < x < b \\ 0, & \text{otherwise} \end{cases}$$



Remark. We have already seen an example of the uniform distribution

$$f_X(x) = 1, \quad 0 < x < 1$$

We can denote this by $X \sim \text{Unif}(0, 1)$. We have also computed the following quantities:

- cdf: $F_X(x) = x, \quad 0 < x < 1$
- expected value: $\text{E}(X) = \frac{1}{2}$
- variance: $\text{Var}(X) = \frac{1}{12}$

Example 0.1. Suppose a bus arrives at a stop uniformly random between noon and 12:15pm, and you arrive at the bus stop exactly at noon. What is the probability that you will wait

(1) no more than 5 minutes or

(2) between 5 and 10 minutes, or

(3) more than 10 minutes?

Theorem 0.1. If $X \sim \text{Unif}(a, b)$, then its cdf is

$$F(x) = \frac{x-a}{b-a}, \quad a < x < b.$$

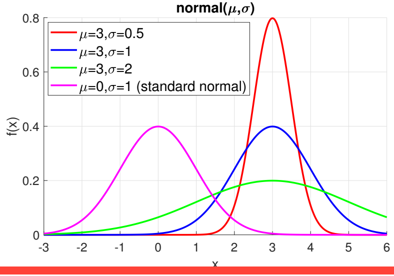
and the mean and variance are

$$\mathrm{E}(X) = \frac{a+b}{2}, \quad \mathrm{Var}(X) = \frac{(b-a)^2}{12}$$

The Normal distribution

Def 0.2 $(X \sim N(\mu, \sigma^2))$**.** We say that a continuous random variable X has a **normal** distribution with parameters μ, σ if it has the following pdf:

$$f(x; \mu, \sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad -\infty < x < \infty$$



(1) The normal curves are all symmetric, unimodal, and bell-shaped;

(2) $\mathrm{E}(X) = \mu$ and $\mathrm{Var}(X) = \sigma^2$;

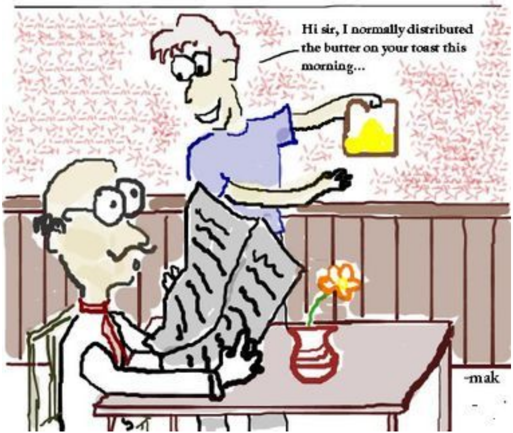
(3) $N(0, 1)$ is called the standard normal distribution:

$f(x; 0, 1) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}, \quad x \in \mathbb{R}$

The Normal distributions are fundamental in probability and statistics:

- Empirically, measurements of a large population often have normal distributions, such as
  - repeated measurements of the same object,
  - heights of a large population, and
  - test scores of a large class.
- Mathematically, one can show that the sums of many independent random variables (individually not necessarily normally distributed) have approximate normal distributions (this result is called the Central Limit Theorem)

Why staticians don't make it as waiters...



Bad news – cdfs of normal distributions do not have explicit formulas: For any given point x_0,

$$F(x_0;\mu,\sigma) = P(X < x_0) = \int_{-\infty}^{x_0} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}\,\mathrm{d}x$$

Through a change of variable

$$z = \frac{x-\mu}{\sigma} \quad \left(\text{and a corresponding change of limit } z_0 = \frac{x_0-\mu}{\sigma}\right)$$

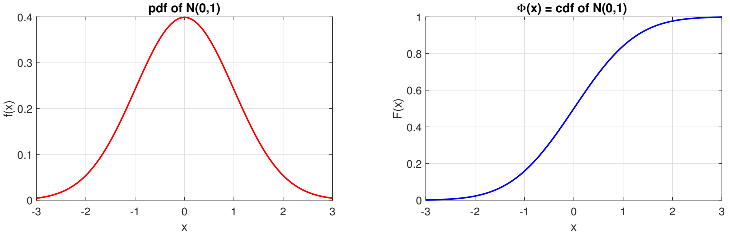
we can obtain that

$$F(x_0;\mu,\sigma) = \int_{-\infty}^{z_0} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}\,\mathrm{d}z = F(z_0;0,1).$$

Good news – All cdf calculations for normal distributions can be reduced to similar calculations for the standard normal.

Displaying the cdf of $N(0,1)$

The cdf of standard normal $\Phi(x) \equiv F(x;0,1)$ can be numerically calculated through a computer:



and displayed in a (huge) table, called standard normal table (linked on `http://www.sjsu.edu/faculty/guangliang.chen/Math161a.html`).

Table entry for z is the area under the standard Normal curve to the left of z.

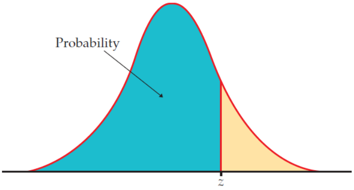


TABLE A

Standard Normal probabilities (continued)

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
0.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
0.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
0.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
0.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
0.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015

Example 0.2. Suppose $Z \sim N(0,1)$. Find

- $P(Z < 0) =$
- $P(Z < -1.3) =$
- $P(Z > 1.3) =$
- $P(-2.5 < Z < 1.5) =$
- $P(-1 < Z < 1) = .6826$
- $P(-2 < Z < 2) = .9544$
- $P(-3 < Z < 3) = .9974$

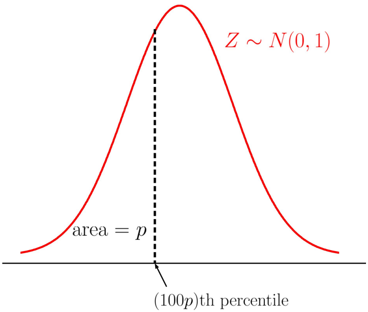
Percentiles

Def 0.3. For any $0 < p < 1$, we define the $(100p)$th **percentile** of the standard normal random variable Z as the cutoff z such that

$$p = P(Z < z) = \Phi(z).$$

Alternatively, we may write

$$z = \Phi^{-1}(p).$$



Example 0.3. Find the 25th (first quartile), 50th (median), 75th (third quartile) percentiles of $Z \sim N(0, 1)$.

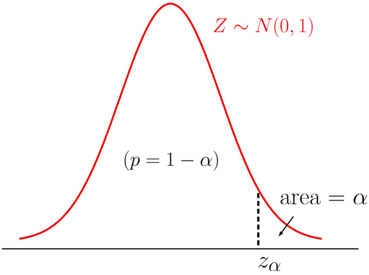
Critical values

Def 0.4. For $0 < \alpha < 1$, we define the z_α **critical value** as

$$P(Z > z_\alpha) = \alpha.$$

Remark. z_α is also the $100(1-\alpha)$th percentile:

$$P(Z < z_\alpha) = 1 - \alpha.$$



Example 0.4. Find z_α for $\alpha = .01, .05, .1$

Standardization

Def 0.5. Let X be a random variable with mean μ and standard deviation σ). Its standardized form is defined as

$$Z = \frac{X - \mu}{\sigma}.$$

Remark. Standardized random variables always have zero mean and unit variance:

$$\mathrm{E}(Z) = \mathrm{E}\left[\frac{1}{\sigma}(X - \mu)\right] = \frac{1}{\sigma}(\mathrm{E}(X) - \mu) = 0$$
$$\mathrm{Var}(Z) = \frac{1}{\sigma^2}\mathrm{Var}(X) = 1.$$

Proposition 0.2. If $X \sim N(\mu, \sigma^2)$, then

$$Z = \frac{X - \mu}{\sigma} \sim N(0,1).$$

Correspondingly,

$$F_X(x; \mu, \sigma) = P(X \leq x) = P\left(Z \leq \frac{x-\mu}{\sigma}\right) = \Phi\left(\frac{x-\mu}{\sigma}\right).$$

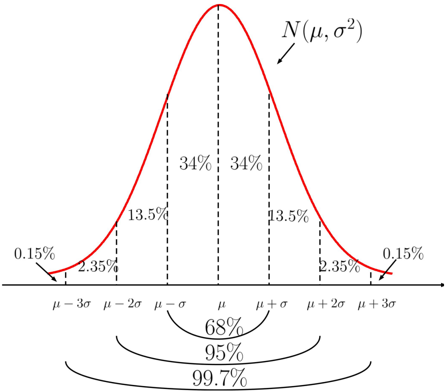
Remark. The normality part of the theorem follows from the fact that any linear transformation of a normal random variable is still normal.

Example 0.5. Suppose $X \sim N(5, 3^2)$. Verify that

$$P(X < -1) = 0.0228$$
$$P(X > 4.1) = 0.6179$$
$$P(2 < X < 5.3) = 0.3811$$

Example 0.6. Suppose $X \sim N(5, 3^2)$. Find the 90th percentile.

The 68-95-99.7 rule

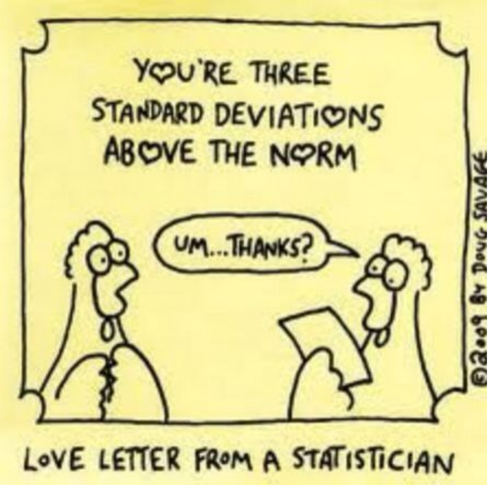


Interpretation: Let $X \sim N(\mu,\sigma^2)$. Then the probabilities of X staying within one/two/three standard deviation around the center are roughly $68\%, 95\%, 99.7\%$, respectively:

$$P(\mu-\sigma < X < \mu+\sigma) \approx .68$$

$$P(\mu-2\sigma < X < \mu+2\sigma) \approx .95$$

$$P(\mu-3\sigma < X < \mu+3\sigma) \approx .997$$


YOU'RE THREE
STANDARD DEVIATIONS
ABOVE THE NORM
UM...THANKS?
©2009 BY DOUG SAVAGE
LOVE LETTER FROM A STATISTICIAN

Normal approximation to binomial

Theorem 0.3. Let $X \sim B(n,p)$. Then for large n such that

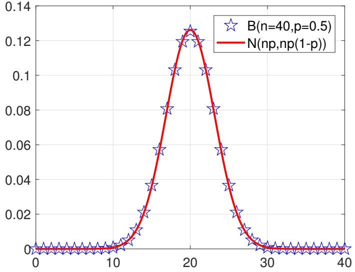
$$np \geq 10,\ n(1-p) \geq 10,$$

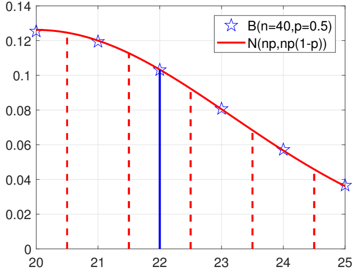
we have

$$X \overset{\text{approx}}{\sim} N(\mu = np, \sigma^2 = np(1-p)),$$

or equivalently,

$$\frac{X - np}{\sqrt{np(1-p)}} \overset{\text{approx}}{\sim} N(0,1).$$





How to make sense of the theorem (we use $x = 22$ as an example):

$$\underbrace{P(X = 22)}_{B(n=40,p=0.5)} \approx \underbrace{P(21.5 < X < 22.5)}_{N(np,np(1-p))}$$

Example 0.7. Use the normal approximation to find the probability of getting exactly 22 heads when tossing a fair coin 40 times.
Answer: Binomial 0.1031, Normal 0.1044

Example 0.8 (Cont'd). What about no more than 22 heads?

Answer: Binomial 0.7852*, Normal approximation* 0.7357*, and Normal+continuity correction* 0.7852

Remark. In the preceding example, the normal approximation to binomial (with continuity correction) works in the following ways:

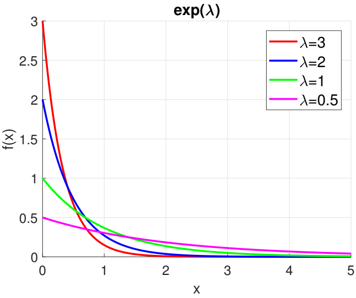
$$P(X = 22) \approx P(21.5 < X < 22.5)$$
$$P(X \leq 22) \approx P(X < 22.5)$$
$$P(X < 22) = P(X \leq 21) \approx P(X < 21.5)$$
$$P(X \geq 22) \approx P(X > 21.5)$$
$$P(X > 22) = P(X \geq 23) \approx P(X > 22.5)$$

The Exponential distribution

Exponential distributions are very useful for modeling the waiting time for a rare event, such as the arrival of a hurricane and the breakdown of an electronic device such as light bulb.

Def 0.6 ($X \sim \text{Exp}(\lambda)$). A continuous random variable X is said to have an **exponential** distribution with parameter λ if its pdf has the following form

$$f(x) = \lambda e^{-\lambda x}, \quad x > 0$$

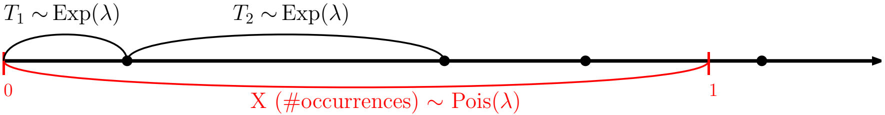


To understand what the parameter λ represents, we first need to find the expected value of $X \sim \mathrm{Exp}(\lambda)$.

Theorem 0.4. If $X \sim \mathrm{Exp}(\lambda)$, then

$$\mathrm{E}(X) = \frac{1}{\lambda}, \quad \mathrm{Var}(X) = \frac{1}{\lambda^2}.$$

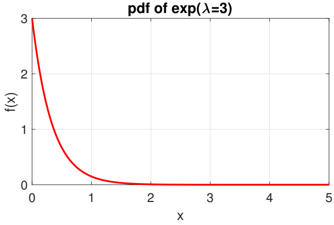
Remark. The preceding theorem indicates that $\frac{1}{\lambda}$ is the mean waiting time for a rare event to occur and thus λ is the rate at which the event occurs (and it is the same parameter lambda of the Poisson distribution).

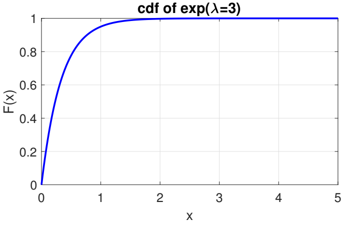


Example 0.9. Suppose that the life time of a certain brand of light bulbs is exponentially distributed with an average of $1,000$ hours. What is the probability that a new light bulb can exceed this amount of time? What about between $1,000$ and $2,000$ hours?

Proposition 0.5. Let $X \sim \mathrm{Exp}(\lambda)$. Then the cdf of X is

$$F_X(x) = 1 - e^{-\lambda x}, \quad x > 0$$





The complementary cdf function

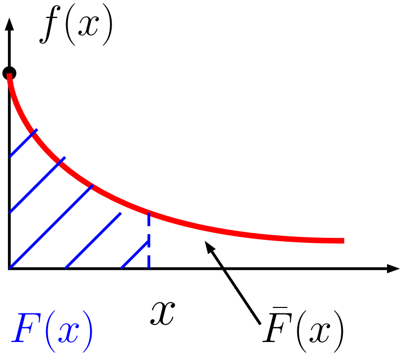
Def 0.7. The **complementary cdf** of a random variable X is defined as

$$\bar{F}(x) = P(X > x) = 1 - F(x)$$

Remark. If $X \sim \mathrm{Exp}(\lambda)$, then

$$\bar{F}(x) = e^{-\lambda x}, \quad x > 0.$$

It can be thought of as the probability of lasting longer than x hours for a light bulb.



Theorem 0.6 (The memoryless property). If $X \sim \mathrm{Exp}(\lambda)$, then

$$P(X > t_0 + t \mid X > t_0) = P(X > t), \quad \text{for any } t_0, t > 0.$$

Interpretation (in the setting of light bulbs):

- $P(X > t)$: probability that a new light bulb can exceed t hours
- $P(X > t_0 + t \mid X > t_0)$: probability that a light bulb can last for t more hours given that it has worked for t_0 hours.

Remark. The exponential distribution is the only continuous distribution that has the memoryless property.

Example 0.10. Jones figures that the total number of thousands of miles that an auto can be driven before it would need to be junked is an exponential random variable with parameter $\lambda = 1/20$. Smith has a used car that he claims has been driven only 10,000 miles. If Jones purchases the car, what is the probability that she would get at least 20,000 additional miles out of it?

Example 0.11 (Cont'd)**.** Repeat under the assumption that the lifetime mileage of the car is not exponentially distributed but rather is (in thousands of miles) uniformly distributed over (0, 40).

The Gamma distribution

The Gamma distribution is defined based on the template function

$$g(x) = x^{\alpha-1}e^{-x}, \quad x > 0,$$

which has a peak at $x = \alpha - 1$ when $\alpha > 1$ and a long right tail:

$$g'(x) = x^{\alpha-2}e^{-x}(\alpha - 1 - x)$$

In order to use $g(x)$ to produce a distribution, we need to normalize it carefully:

$$1 = \int_0^\infty Cx^{\alpha-1}e^{-x}\,\mathrm{d}x = C \cdot \underbrace{\int_0^\infty x^{\alpha-1}e^{-x}\,\mathrm{d}x}_{\Gamma(\alpha)} \longrightarrow C = \frac{1}{\Gamma(\alpha)}.$$

The Gamma function

Def 0.8. The **Gamma function** is a function $\Gamma : (0,\infty) \mapsto (0,\infty)$ with

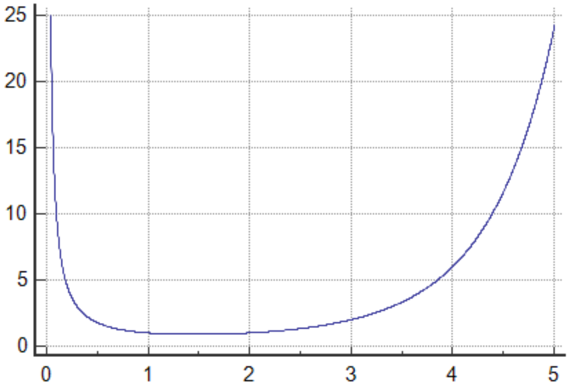
$$\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} e^{-x}\, \mathrm{d}x, \quad \alpha > 0$$

(The Gamma function can be seen as a way to generalize factorials from integers to non-integers, e.g., 2.4!)

Properties:

- $\Gamma(1) = 1$
- *For any* $\alpha > 0$, $\Gamma(\alpha+1) = \alpha \cdot \Gamma(\alpha)$
- *For any positive integer* n, $\Gamma(n) = (n-1)!$
- $\Gamma(\frac{1}{2}) = \sqrt{\pi}$

Graph of the Gamma function



https://www.medcalc.org/manual/gamma_function.php

We introduce a second parameter (β) to make the Gamma distribution more flexible.

From

$$1 = \int_0^\infty \frac{1}{\Gamma(\alpha)} x^{\alpha-1} e^{-x} \,\mathrm{d}x,$$

by letting $x = y/\beta$ for some $\beta > 0$, we have

$$\begin{aligned} 1 &= \int_0^\infty \frac{1}{\Gamma(\alpha)} \left(\frac{y}{\beta}\right)^{\alpha-1} e^{-y/\beta} \, \frac{1}{\beta} \,\mathrm{d}y \\ &= \int_0^\infty \underbrace{\frac{1}{\beta^\alpha \,\Gamma(\alpha)} y^{\alpha-1} e^{-y/\beta}}_{\text{two-parameter Gamma density}} \,\mathrm{d}y \end{aligned}$$

The two-parameter Gamma distribution

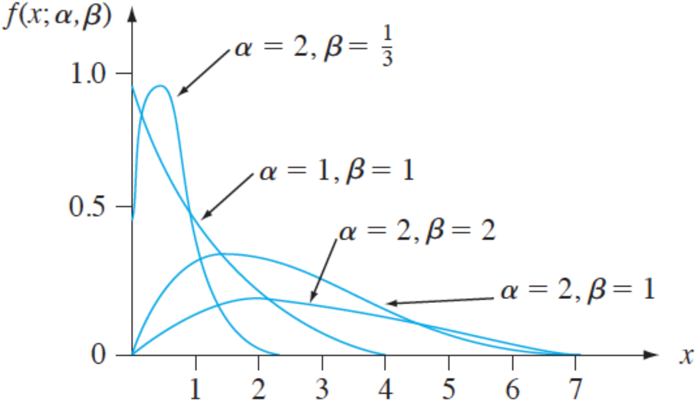
Def 0.9 ($X \sim \text{Gamma}(\alpha, \beta)$). A random variable X is said to have a (two-parameter) **Gamma distribution** with parameters α, β if it has a pdf of the form

$$f(x; \alpha, \beta) = \frac{1}{\beta^\alpha \, \Gamma(\alpha)} x^{\alpha-1} e^{-x/\beta}, \quad x > 0$$

Remark. If $\alpha = 1$, then Gamma(α, β) reduces to Exp($\lambda = 1/\beta$).

Theorem 0.7. If $X \sim \text{Gamma}(\alpha, \beta)$, then

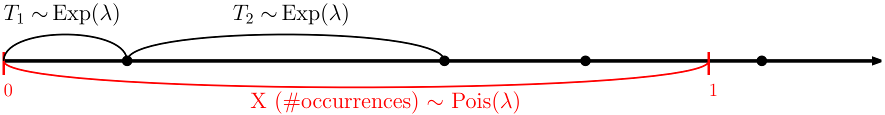
$$\text{E}(X) = \alpha\beta, \quad \text{Var}(X) = \alpha\beta^2$$



(β is a scale parameter)

Application of the Gamma distribution

Consider the experiment of counting the occurrences of a rare event (such as hurricane) that occurs with rate λ:



It is already known that

- The total number of occurrences of the event in a unit interval of time has a Poisson distribution: $X \sim \mathrm{Pois}(\lambda)$;

- The separate waiting times for different occurrences of the event, $T_1, T_2, \ldots$, are iid Exp(λ).

It turns out that the total waiting time for n occurrences of the event has a Gamma distribution:

$$T = T_1 + \cdots + T_n \sim \text{Gamma}(\alpha = n, \beta = 1/\lambda)$$

This implies that

$$\text{E}(T) = \text{E}(T_1) + \cdots + \text{E}(T_n) = n \cdot \frac{1}{\lambda} = \frac{n}{\lambda}$$
$$\text{Var}(T) = \text{Var}(T_1) + \cdots + \text{Var}(T_n) = n \cdot \frac{1}{\lambda^2} = \frac{n}{\lambda^2}$$

The chi-squared distribution

Another special case of the Gamma distribution is the **chi-squared distribution with parameter** k, denoted as $\chi^2(k)$ and sometimes also χ^2_k:

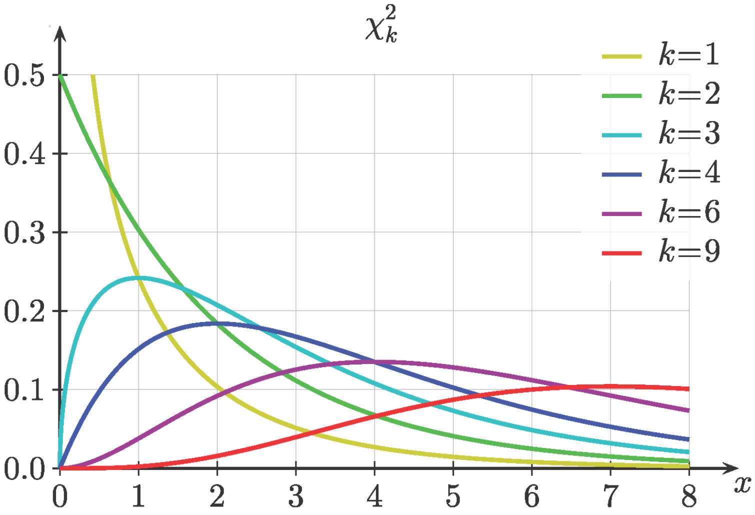
$$\text{Gamma}(\alpha = \frac{k}{2},\ \beta = 2) = \chi^2(k) \quad \longleftarrow k \text{ is called \#degrees of freedom}$$

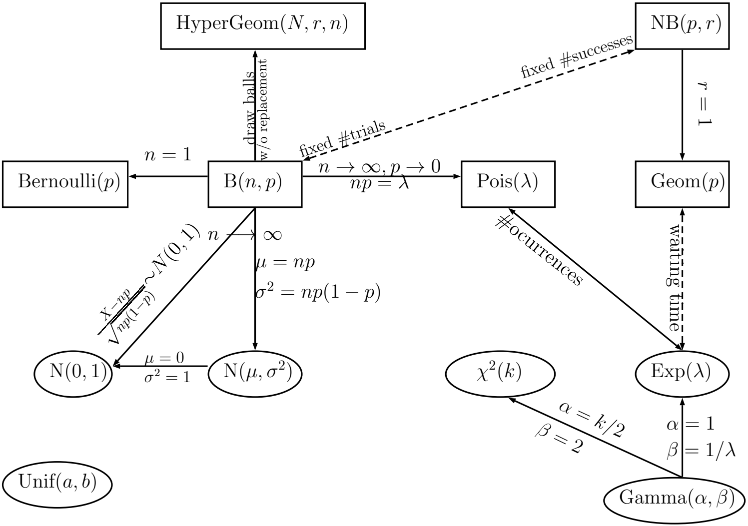
It is also the distribution of $X = Z_1^2 + \cdots + Z_k^2$ where $Z_1, \ldots, Z_k \overset{iid}{\sim} N(0,1)$.

The pdf of the $\chi^2(k)$ distribution is the following:

$$f(x) = \frac{1}{2^{k/2}\,\Gamma(k/2)} x^{(k/2)-1} e^{-x/2}, \quad x > 0$$

Its mean and variance are $\mathrm{E}(X) = k$ and $\mathrm{Var}(X) = 2k$.


χ_k^2
0.5
0.4
0.3
0.2
0.1
0.0
0
1
2
3
4
5
6
7
8
x
$k=1$
$k=2$
$k=3$
$k=4$
$k=6$
$k=9$


HyperGeom(N, r, n)
NB(p, r)
draw balls w/o replacement
fixed #successes
fixed #trials
$r = 1$
$n = 1$
Bernoulli(p)
B(n, p)
$n \to \infty, p \to 0$
$np = \lambda$
Pois(λ)
Geom(p)
$\frac{X-np}{\sqrt{np(1-p)}} \sim N(0,1)$
$n \to \infty$
$\mu = np$
$\sigma^2 = np(1-p)$
#ocurrences
waiting time
N$(0,1)$
$\mu = 0$
$\sigma^2 = 1$
N(μ, σ^2)
$\chi^2(k)$
Exp(λ)
$\alpha = k/2$
$\beta = 2$
$\alpha = 1$
$\beta = 1/\lambda$
Unif(a, b)
Gamma(α, β)