VE401 RC Week3

Wang Yangyang

UM-SJTU JI

2022 Spring

Outline

Outline

Random Variables

Assumption

- We distinguish between
  - ***discrete random variables***, defined as having a countable range in $\mathbb{R}$
  - ***continuous random variables***, defined as having range equal to $\mathbb{R}$.
- We assume that a random variable comes with a ***probability density*** function that allows the calculation of probabilities directly, without recourse to the probability space.

Discrete Random Variables and PDF

Definition

Let S be a sample space and Ω a countable subset of $\mathbb{R}$. A *discrete random variable* is a map

$$X : S \to \Omega$$

together with a function

$$f_X : \Omega \to \mathbb{R}$$

having the properties that

- $f_X(x) \geqslant 0$ for all $x \in \Omega$ and
- $\sum_{x \in \Omega} f_X(x) = 1$.

The function f_X is called the *probability density function* or *probability distribution* of X.
A *random variable* is given by the pair (X, f_X).

Cumulative Density Function

Definition

The cumulative distribution function of a random variable is defined as

$$F_X : \mathbb{R} \to \mathbb{R}, \quad F_X(x) := P[X \leqslant x]$$

For a discrete random variable,

$$F_X(x) = \sum_{y \leqslant x} f_X(y)$$

Outline

Expectation

Definition

Let (X, f_X) be a discrete random variable. Then the expected value or expectation of X is

$$\mathrm{E}[X] := \sum_{x \in \Omega} x \cdot f_X(x)$$

provided that the sum on the right converges absolutely.

We often write μ_X or simply μ for the expectation.

Lemma

Let (X, f_X) be a discrete random variable and $\varphi : \Omega \to \mathbb{R}$ some function. Then the expected value of $\varphi \circ X$ is

$$\mathrm{E}[\varphi \circ X] = \sum_{x \in \Omega} \varphi(x) \cdot f_X(x)$$

provided that the sum (series) on the right converges absolutely.

Expectation

Properties

- Linearity
  The expected value operator (or expectation operator) $\mathrm{E}[\cdot]$ is linear in the sense that, for any random variables X and Y, and a constant a,
  $$\mathrm{E}[X+Y] = \mathrm{E}[X] + \mathrm{E}[Y]$$
  $$\mathrm{E}[aX] = a\mathrm{E}[X]$$
  whenever the right-hand side is well-defined.
  Symbolically, for N random variables X_i and constants $a_i (1 \leqslant i \leqslant N)$, we have $\mathrm{E}\left[\sum_{i=1}^{N} a_i X_i\right] = \sum_{i=1}^{N} a_i \mathrm{E}\left[X_i\right]$
- Non-Multiplicity
  If $X_1, \ldots, X_n$ are ***independent*** random variables with finite expectations then $\mathrm{E}\left[\prod_{i=1}^{n} X_i\right] = \prod_{i=1}^{n} \mathrm{E}\left[X_i\right]$.
  However, if $X_1, \ldots, X_n$ are ***not independent***, then $\mathrm{E}\left[\prod_{i=1}^{n} X_i\right] \neq \prod_{i=1}^{n} \mathrm{E}\left[X_i\right]$.

Variance

Definition

The variance is defined by

$$\sigma_X^2 = \mathsf{Var}[X] := \mathrm{E}\left[(X - \mathrm{E}[X])^2\right]$$

which is defined as long as the right-hand side exists.
The standard deviation is $\sigma_X = \sqrt{\mathsf{Var}[X]}$.

While expectation can be seen as a measure of location, variance can be seen as a measure of dispersion.

Useful Formula

$$\begin{aligned}\mathsf{Var}[X] &= \mathrm{E}\left[(X - \mathrm{E}[X])^2\right] \\ &= \mathrm{E}\left[X^2 - 2\mathrm{E}[X]\cdot X + \mathrm{E}[X]^2\right] \\ &= \mathrm{E}\left[X^2\right] - \mathrm{E}[X]^2\end{aligned}$$

Variance

Properties

- Variance is non-negative. $\mathrm{Var}(X) \geqslant 0$
- Variance of a constant is zero. $\mathrm{Var}(a) = 0$
- Variance is invariant with respect to constant added to all values of the variable. $\mathrm{Var}(X + a) = \mathrm{Var}(X)$
- If all values are scaled by a constant, the variance is scaled by the square of that constant. $\mathrm{Var}(aX) = a^2 \mathrm{Var}(X)$
- The variance of a sum of two ***independent*** random variables is given by

$$\mathrm{Var}(aX + bY) = a^2 \mathrm{Var}(X) + b^2 \mathrm{Var}(Y)$$
$$\mathrm{Var}(aX - bY) = a^2 \mathrm{Var}(X) + b^2 \mathrm{Var}(Y)$$

(If not independent, then covariance should be considered.)

Variance

Example

Prove $\text{Var}[X+Y] = \text{Var}[X-Y] = \text{Var}[X] + \text{Var}[Y]$ for two ***independent*** random variables X, Y.

Variance

Example

Prove $\mathrm{Var}[X+Y] = \mathrm{Var}[X-Y] = \mathrm{Var}[X] + \mathrm{Var}[Y]$ for two ***independent*** random variables X, Y.

Solution

$$\begin{aligned}\mathrm{Var}[X+Y] &= \mathrm{E}\left[(X+Y-(\mu_X+\mu_Y))^2\right] \\ &= \mathrm{E}\left[(X-\mu_X)^2\right] + \mathrm{E}\left[(Y-\mu_Y)^2\right] \\ &\quad + 2\mathrm{E}\left[(X-\mu_X)(Y-\mu_Y)\right]\end{aligned}$$

If X, Y are independent,
$\mathrm{E}\left[(X-\mu_X)(Y-\mu_Y)\right] = \mathrm{E}[X-\mu_X]\mathrm{E}[Y-\mu_Y] = 0 \times 0 = 0.$

Ordinary and Central Moments

Definition

The n^{th} (ordinary) ***moments*** of a random variable X is given by

$$\mathrm{E}\left[X^n\right], \quad n \in \mathbb{N}.$$

The n^{th} ***central moments*** of X is given by

$$\mathrm{E}\left[\left(\frac{X-\mu}{\sigma}\right)^n\right], \quad \text{where } n = 3, 4, 5, \ldots$$

Moment-Generating Function

Definition

Let (X, f_X) be a random variable and such that the sequence of moments $E\left[X^n\right], n \in \mathbb{N}$, exists. If the power series

$$m_X(t) := \sum_{k=0}^{\infty} \frac{\mathrm{E}\left[X^k\right]}{k!} t^k$$

has radius of convergence $\varepsilon > 0$, the thereby defined function

$$m_X(t) : (-\varepsilon, \varepsilon) \to \mathbb{R}$$

is called the ***moment-generating function*** for X.

Moment-Generating Function

Theorem

Let $\varepsilon > 0$ be given such that $\mathrm{E}\left[e^{tX}\right]$ exists and has a power series expansion in t that converges for $|t| < \varepsilon$. Then the ***moment-generating function*** exists and

$$m_X(t) = \mathrm{E}\left[e^{tX}\right] \qquad \text{for } |t| < \varepsilon$$

Furthermore,

$$\mathrm{E}\left[X^k\right] = \left.\frac{\mathrm{d}^k m_X(t)}{\mathrm{d}t^k}\right|_{t=0}$$

We can hence calculate the moments of X by differentiating the moment-generating function.

Moment-Generating Function

Properties

- X is a random variable and $Y = aX + b$, $a, b \in \mathbb{R}$, then for every t such that $m_X(at)$ is finite,

$$m_Y(t) = e^{bt} m_X(at).$$

- Suppose $X_1, \ldots, X_n$ are n independent random variables, then for every value that $m_{X_i}(t)$ is finite for all $i = 1, \ldots, n$,

$$m_X(t) = \prod_{i=1}^{n} m_{X_i}(t), \quad X = X_1 + \cdots + X_n.$$

Moment-Generating Function

Example

Suppose that X is a random variable with the moment-generating function

$$m_X : \mathbb{R} \to \mathbb{R}, \quad m_X(t) = e^{t^2+3t}$$

Find the mean and variance of X.

Moment-Generating Function

Example

Suppose that X is a random variable with the moment-generating function

$$m_X : \mathbb{R} \to \mathbb{R}, \quad m_X(t) = e^{t^2+3t}$$

Find the mean and variance of X.

Solution

We calculate

$$m_X'(t) = (2t+3)e^{t^2+3t}, \quad m_X''(t) = (2t+3)^2 e^{t^2+3t} + 2e^{t^2+3t}$$

Therefore,

$$\mu = m_X'(0) = 3, \quad \sigma^2 = \mathrm{E}\left[X^2\right] - \mathrm{E}[X]^2 = m_X''(0) - \mu^2 = 2$$

Outline

Bernoulli Distribution

Definition

A random variable (X, f_X) has a ***Bernoulli distribution*** with parameter $p, 0 < p < 1$ if the probability density function is defined by

$$f_X : \{0, 1\} \to \mathbb{R}, \quad f_X(x) = \begin{cases} 1 - p, & \text{if } x = 0 \\ p, & \text{if } x = 1 \end{cases}$$

Interpretation

Describe the probability of success $f_X(1)$ or failure $f_X(0)$ of a trial, given the probability of success is p.

Bernoulli Distribution

Properties

- Mean

$$\mathrm{E}[X] = 0 \cdot (1-p) + 1 \cdot p = p$$

- Variance

$$\mathrm{Var}[X] = \mathrm{E}\left[X^2\right] - \mathrm{E}[X]^2 = p - p^2 = p(1-p)$$

- M.G.F.

$$m_X : \mathbb{R} \to \mathbb{R}, \quad m_X(t) = 1 - p + pe^t$$

Binomial Distribution

Definition

A random variable (X, f_X) has a *Binomial distribution* with parameter $n \in \mathbb{N} \backslash \{0\}$ and $p, 0 < p < 1$ if it has probability density function

$$f_X : \{0, \ldots, n\} \to \mathbb{R}, \quad f_X(x) = \binom{n}{x} p^x (1-p)^{n-x}$$

Interpretation

$f_X(x)$ is the probability of obtaining x successes in n independent and identical Bernoulli trials with parameter p.

Binomial Distribution

Properties

- Mean

$$\mathrm{E}[X] = \sum_{i=1}^{n} \mathrm{E}\left[X_i\right] = np$$

- Variance

$$\mathrm{Var}[X] = \sum_{i=1}^{n} \mathrm{Var}\left[X_i\right] = np(1-p)$$

- M.G.F.

$$m_X : \mathbb{R} \to \mathbb{R}, \quad m_X(t) = \mathrm{E}\left[e^{tX}\right] = \prod_{i=1}^{n} \mathrm{E}\left[e^{tX_i}\right] = \left(1 - p + pe^t\right)^n$$

Outline

Geometric Distribution

Definition

A random variable (X, f_X) has ***Geometric distribution*** with parameter $p, 0 < p < 1$ if the probability density function is given by

$$f_X : \mathbb{N}\backslash\{0\} \to \mathbb{R}, \quad f_X(x) = (1-p)^{x-1}p.$$

Interpretation

$f_X(x)$ is the probability of x failures before the first success in the Bernoulli trials, given the probability of success for each trial is p.

Geometric Distribution

Properties

Let $q = 1 - p$,

- Mean
$$\mathrm{E}[X] = \frac{1}{p}$$
- Variance
$$\mathrm{Var}[X] = \frac{q}{p^2}$$
- M.G.F.
$$m_X : (-\infty, -\ln q) \to \mathbb{R}$$
$$m_X(t) = \sum_{k=1}^{\infty} e^{tk} p q^{k-1} = p e^t \sum_{k=1}^{\infty} (e^t q)^{k-1} = \frac{p e^t}{1 - q e^t}$$

The geometric series only converges if $e^t q < 1$.

Pascal Distribution

Definition

A random variable (X, f_X) has the ***Pascal distribution*** with parameters $p, 0 < p < 1$ and $r \in \mathbb{N} \backslash \{0\}$ if the probability density function is given by

$$f_X : \{r, r+1, \ldots\} \to \mathbb{R}, \quad f_X(x) = \binom{x-1}{r-1} p^r (1-p)^{x-r}$$

Interpretation

$f_X(x)$ is the probability of obtaining the r-th success in the x-th Bernoulli trial, given the probability of success for each trial is p.

Pascal Distribution

Properties

Let $q = 1 - p$,

- Mean
$$\mathrm{E}[X] = \frac{r}{p}$$
- Variance
$$\mathrm{Var}[X] = \frac{rq}{p^2}$$
- M.G.F.
$$m_X : (-\infty, -\ln q) \to \mathbb{R}, \quad m_X(t) = \frac{(pe^t)^r}{(1 - qe^t)^r}$$

Negative Binomial Distribution

Definition

A random variable (X, f_X) has the *Negative Binomial distribution* with parameters r and p if the probability density function is given by

$$f_X : \mathbb{N} \to \mathbb{R}, \quad f_X(x) = \binom{x+r-1}{r-1} p^r (1-p)^x$$

Interpretation

$f_X(x)$ is the probability of x failures before first obtaining r successes in Bernoulli trials, given the probability for each success is p.

Negative Binomial Distribution

Properties

Let $q = 1 - p$,

- Mean

$$\mathrm{E}[X] = \frac{rp}{q}$$

- Variance

$$\mathrm{Var}[X] = \frac{rp}{q^2}$$

- M.G.F.

$$m_X : (-\infty, -\ln q) \to \mathbb{R}, \quad m_X(t) = \frac{p^r}{(1 - qe^t)^r}$$

Outline

Poisson Distribution

Definition

A random variable X has the *Poisson distribution* with parameter $k > 0$ if probability density function is given by

$$f_X : \mathbb{N} \to \mathbb{R}, \quad f_X(x) = \frac{k^x e^{-k}}{x!}$$

Interpretation

$f_X(x)$ is the probability of x arrivals in the time interval $[0, t]$ with arrival rate $\lambda > 0$, and $k = \lambda t$.
"...which describes the occurrence of events that occur at *constant rate* and *continuous environment*."

Poisson Distribution

Interpretation

- *Continuous environment*. Not limited to time intervals, but also subregions of two- or three-dimensional regions or sublengths of a linear distance, and any regions that can be divided into arbitrarily small pieces.
- *Constant rate*. The probability of an occurrence during each very short interval (region) must be approximately proportional to the length (area, volume) of that interval (region).

Example

Poisson process can be used to model

- the number of particles that strike a certain target at a constant rate in a particular period;
- the number of oocysts that occur in a water supply system given constant rate of occurrence per liter...

Poisson Distribution

Properties

- Mean

$$\mathrm{E}[X] = k$$

- Variance

$$\mathrm{Var}[X] = k$$

- M.G.F.

$$m_X : \mathbb{R} \to \mathbb{R}, \quad m_X(t) = e^{k(e^t - 1)}$$

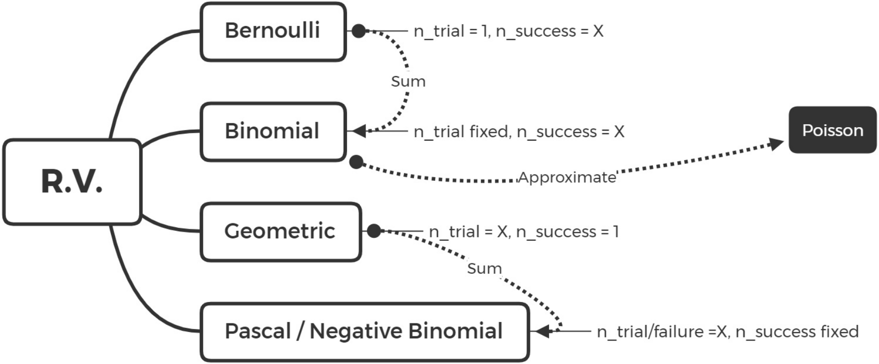
Derivation

$$m_X(t) = \mathrm{E}\left(e^{tX}\right) = \sum_{n=0}^{\infty} P(X = n) e^{tn} = \sum_{n=0}^{\infty} \frac{\lambda^n e^{-\lambda}}{n!} e^{tn}$$

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

Outline

Connections



Summations

- Bernoulli → Binomial. $X_1, \ldots, X_n$ are independent random variables,

$$X_i \sim \text{Bernoulli}(p) \quad \Rightarrow \quad X = X_1 + \cdots + X_n \sim \mathrm{B}(n, p).$$

- Geometric → Pascal/Negative Binomial. $X_1, \ldots, X_r$ are independent random variables,

$$X_i \sim \text{Geom}(p) \quad \Rightarrow \quad X = X_1 + \cdots + X_r \sim \text{Pascal}(r, p).$$

- Poisson → Poisson. $X_1, \ldots, X_n$ are independent random variables,

$$X_i \sim \text{Poisson}\left(k_i\right) \quad \Rightarrow \quad X = X_1 + \cdots + X_n \sim \text{Poisson}(k)$$

where $k = k_1 + \cdots + k_n$.

Summations

- Binomial → Binomial. $X_1, \ldots, X_k$ are independent random variables,

$$X_i \sim B\left(n_i, p\right) \quad \Rightarrow \quad X = X_1 + \cdots + X_k \sim B(n, p),$$

  where $n = n_1 + \cdots + n_k$.
- Negative Binomial → Negative Binomial. $X_1, \ldots, X_n$ are independent random variables,

$$X_i \sim \mathrm{NB}\left(r_i, p\right) \quad \Rightarrow \quad X = X_1 + \cdots + X_n \sim \mathrm{NB}(r, p),$$

  where $r = r_1 + \cdots + r_n$.
- Pascal → Pascal. $X_1, \ldots, X_n$ are independent random variables,

$$X_i \sim \mathrm{Pascal}\left(r_i, p\right) \quad \Rightarrow \quad X = X_1 + \cdots + X_n \sim \mathrm{Pascal}(r, p),$$

  where $r = r_1 + \cdots + r_n$.

Summations

Example

Let X be a discrete random variable following a ***Geometric distribution*** with parameter $p = 1/2$ and let $X_1, \ldots, X_{10}$ be a random sample of size 10 . Calculate the probability that the sample mean is no more than than 1.5, i.e., find $P[\overline{X} \leqslant 1.5]$.

Summations

Example

Let X be a discrete random variable following a ***Geometric distribution*** with parameter $p = 1/2$ and let $X_1, \ldots, X_{10}$ be a random sample of size 10 . Calculate the probability that the sample mean is no more than than 1.5, i.e., find $P[\overline{X} \leqslant 1.5]$.

Solution

We note that $X_1 + \cdots + X_{10}$ follows a ***Pascal distribution*** with $r = 10$ and $p = 1/2$. Then

$$P[\overline{X} < 1.5] = P\left[X_1 + \cdots + X_{10} \leqslant 10 \cdot 1.5\right] = P\left[X_1 + \cdots + X_{10} \leqslant 15\right]$$

$$= \sum_{x=10}^{15} \binom{x-1}{9} \frac{1}{2^x} = \frac{309}{2048} = 0.15$$

How about ***Negative Binomial*** with $r = 10$ and $p = 1/2$?

Closeness Between Poisson Distribution and Binomial Distribution

Theorem

For $n \in \mathbb{N} \backslash \{0\}, 0 < p < 1$, suppose $f(x; n, p)$ denotes the probability density function of ***Binomial distribution*** with parameters n and p, while $f(x; k)$ denotes the probability density function of ***Poisson distribution*** with parameter k. Let $\{p_n\}_{n=1}^{\infty}$ be a sequence of numbers between 0 and 1 such that

$$\lim_{n\to\infty} np_n = k$$

then

$$\lim_{n\to\infty} f(x; n, p_n) = f(x; k), \qquad \text{for all } x = 0, 1, \ldots$$

This means we can approximate the ***Binomial distribution*** with ***Poisson distribution*** when n is large.

Outline

Continuous Random Variable: Def. and PDF

Definition

Let S be a sample space. A *continuous random variable* is a map $X : S \to \mathbb{R}$ together with a function $f_X : \mathbb{R} \to \mathbb{R}$ with the properties that

1. $f_X(x) \geqslant 0$ for all $x \in \mathbb{R}$ and
2. $\int_{-\infty}^{\infty} f_X(x)\mathrm{d}x = 1$.

The integral of f_X is interpreted as the probability that X assumes values X in a given range, i.e.,

$$P[a \leqslant X \leqslant b] = \int_a^b f_X(x)\mathrm{d}x$$

The function f_X is called the *probability density function* of random variable X.

Continuous Random Variable: CDF and Location

Definition

Let (X, f_X) be a continuous random variable. The *cumulative distribution function* for X is defined by $F_X : \mathbb{R} \to \mathbb{R}$,

$$F_X(x) := P[X \leqslant x] = \int_{-\infty}^{x} f_X(y)\mathrm{d}y$$

By the fundamental theorem of calculus, we can obtain the density function from F_X by

$$f_X(x) = F_X'(x)$$

Definition

- The *median* M_X is defined by $P\left[X \leqslant M_X\right] = 0.5$.
- The *mean* is given by $\mathrm{E}[X]$.
- The *mode* x_0, is the location of the maximum of f_X.

Outline

Expectation, Variance and Moments

- Expectation.

$$\mathrm{E}[X] := \int_{\mathbb{R}} x \cdot f_X(x)\mathrm{d}x$$

- Variance.

$$\mathrm{Var}[X] := \mathrm{E}\left[(X - \mathrm{E}[X])^2\right] = \mathrm{E}\left[X^2\right] - \mathrm{E}[X]^2$$

- M.G.F.

$$m_X(t) = \mathrm{E}\left[e^{tX}\right] = \int_{-\infty}^{\infty} e^{tx} f_X(x)\mathrm{d}x$$

Note: All previous properties about expectation, variance and M.G.F. hold for continuous random variables.

Outline

Exponential Distribution

Definition

A continuous random variable (X, f_β) follows ***exponential distribution*** with parameter β if the probability density function is defined by

$$f_\beta(X) = \begin{cases} \beta e^{-\beta x}, & x > 0 \\ 0, & x \leqslant 0 \end{cases}$$

Interpretation

The time between successive arrivals of a Poisson process with rate λ follows exponential distribution with parameter $\beta = \lambda$. (Recall $P[T > t] = e^{-\beta t}$.) ***Note.*** Memoryless property:

$$P[X > x + s \mid X > x] = P[X > s]$$

Exponential Distribution

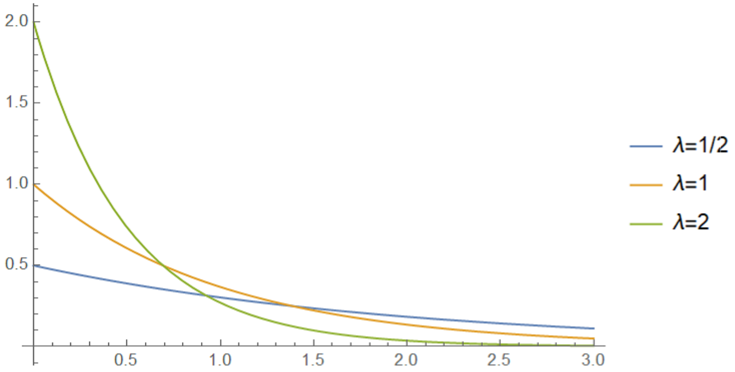
Properties

- Mean.
$$\mathrm{E}[X] = \frac{1}{\beta}$$
- Variance.
$$\mathrm{Var}[X] = \frac{1}{\beta^2}$$
- M.G.F.
$$m_X : (-\infty, \beta) \to \mathbb{R}, \quad m_X(t) = \frac{1}{1 - t/\beta}$$

Exponential Distribution

Plot

```
Plot[Table[PDF[ExponentialDistribution[λ], x], {λ, {1/2, 1, 2}}]
// Evaluate, {x, 0, 3}, PlotRange -> All,
PlotLegends -> {"λ=1/2", "λ=1", "λ=2"}]
```



Exponential Distribution

Example

n light bulbs are burning simultaneously and independently, and the lifetime for each bulb follows the $\text{Exp}(\beta)$.

- What is the distribution of the length of time Y_1 until the first failure in one of the n bulbs?
- What is the distribution of the length of time Y_2 after the first failure until a second bulb fails?

Exponential Distribution

Example

n light bulbs are burning simultaneously and independently, and the lifetime for each bulb follows the $\text{Exp}(\beta)$.

- What is the distribution of the length of time Y_1 until the first failure in one of the n bulbs?
- What is the distribution of the length of time Y_2 after the first failure until a second bulb fails?

Solution

(i) Suppose random variables $X_1, \ldots, X_n$ satisfies that $X_i \sim \text{Exp}(\beta)$, and $Y_1 = \min\{X_1, \ldots, X_n\}$. Then for any $t > 0$,

$$P[Y_1 > t] = P[X_1 > t, \ldots, X_n > t]$$

$$= P[X_1 > t] \times \cdots \times P[X_n > t] = e^{-n\beta t}$$

indicating an exponential distribution with parameter $n\beta$.

Exponential Distribution

Example

n light bulbs are burning simultaneously and independently, and the lifetime for each bulb follows the $\text{Exp}(\beta)$.

- What is the distribution of the length of time Y_1 until the first failure in one of the n bulbs?
- What is the distribution of the length of time Y_2 after the first failure until a second bulb fails?

Solution

(ii) Exponential Distribution is memoryless.

$$P[Y_2 > t] = P[X_2 > t, \ldots, X_n > t]$$

$$= P[X_2 > t] \times \cdots \times P[X_n > t] = e^{-(n-1)\beta t}$$

indicating an exponential distribution with parameter $(n-1)\beta$.

Gamma Distribution

Definition

Let $\alpha, \beta \in \mathbb{R}$, $\alpha, \beta > 0$. A continuous random variable $(X, f_{\alpha,\beta})$ follows a gamma distribution with parameters α and β if the probability density function is given by

$$f_{\alpha,\beta}(x) = \begin{cases} \frac{\beta^{\alpha}}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}, & x > 0 \\ 0, & x \leqslant 0 \end{cases}$$

where $\Gamma(\alpha) = \int_0^{\infty} z^{\alpha-1} e^{-z}\,\mathrm{d}z$, $\alpha > 0$ is the Euler gamma function.
Interpretation.

Interpretation

The time needed for the next r arrivals in a Poisson process with rate λ follows a Gamma distribution with parameters $\alpha = r, \beta = \lambda$.

Gamma Distribution

Properties

- Mean.

$$\mathrm{E}[X] = \frac{\alpha}{\beta}$$

- Variance.

$$\mathrm{Var}[X] = \frac{\alpha}{\beta^2}$$

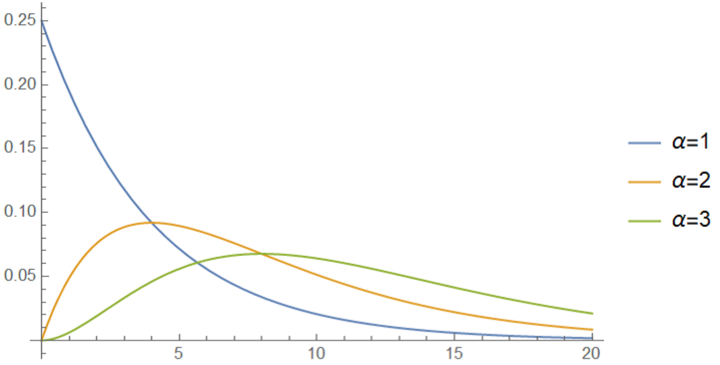
- M.G.F.

$$m_X : (-\infty, \beta) \to \mathbb{R}, \quad m_X(t) = \frac{1}{(1 - t/\beta)^\alpha}$$

Gamma Distribution

Plot

```
Plot[Table[PDF[GammaDistribution[α,4], x], {α, {1, 2, 3}}]
// Evaluate, {x, 0, 20}, PlotRange -> All,
PlotLegends -> {"α=1", "α=2", "α=3"}]
```

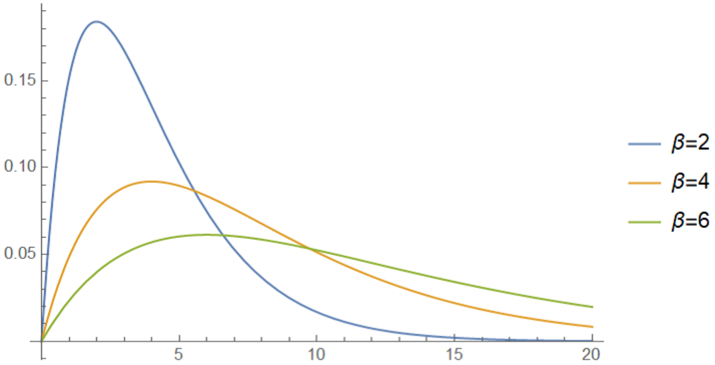


Gamma Distribution

Plot

What's Wrong? – FOR β, MMA IS DIFFERENT FROM THE LECTURE! It's $(1/2, 1/4, 1/6)$ for the lecture version.

```
Plot[Table[PDF[GammaDistribution[2,β], x], {β, {2, 4, 6}}]
// Evaluate, {x, 0, 20}, PlotRange -> All,
PlotLegends -> {"β=2", "β=4", "β=6"}]
```



Chi-Squared Distribution

Definition

Definition. Let $\gamma \in \mathbb{N}$. A continuous random variable $\left(X_\gamma^2, f_X\right)$ follows a chi-squared distribution with γ degrees of freedom if the probability density function is given by

$$f_\gamma(x) = \begin{cases} \frac{1}{2^{\gamma/2}\Gamma(\gamma/2)} x^{\gamma/2-1} e^{-x/2}, & x > 0 \\ 0, & x \leqslant 0 \end{cases}$$

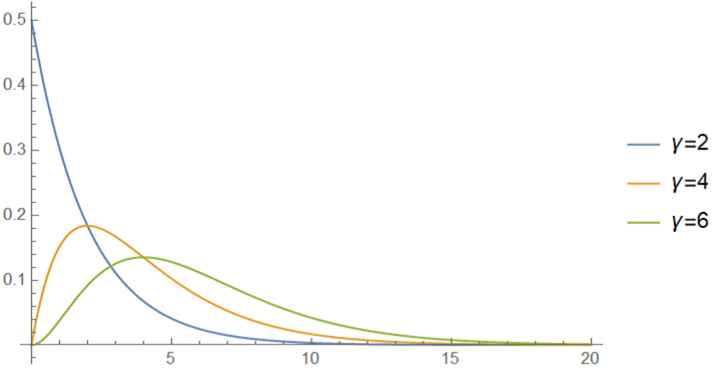
It is a gamma distribution with $\alpha = \gamma/2, \beta = 1/2$. Therefore,

$$\mathrm{E}\left[X_\gamma^2\right] = \gamma, \quad \mathrm{Var}\left[X_\gamma^2\right] = 2\gamma$$

Chi-Squared Distribution

Plot

```
Plot[Table[PDF[ChiSquareDistribution[γ], x], {γ, {2, 4, 6}}]
// Evaluate, {x, 0, 20}, PlotRange -> All,
PlotLegends -> {"γ=2", "γ=4", "γ=6"}]
```



Outline

1.Integral to Distribution

Exercise

For $0 < p < 1$ and $n = 2, 3, \ldots$, determine the value of

$$\sum_{x=2}^{n} x(x-1) \binom{n}{x} p^x (1-p)^{n-x}$$

2.Poisson Approximation to the Binomial

Discussion

Prove that we can approximate the ***Binomial distribution*** with ***Poisson distribution*** when n is large (Have Fun!).

Exercise

A factory produces drill bits. It is known that 2% of the drill bits are defective. The drill bits are shipped in packages of 100 . Use the Poisson approximation to the binomial distribution to answer the following questions.

1. What is the probability that there are no more than three defective drill bits in a package?
2. How many drill bits must a package contain so that with probability greater than 0.9 there are at least 100 non-defective drill bits in the package?

3.Relationship Between Poisson and Exponential

Discussion

Prove that the time between successive arrivals of a *Poisson-distributed* random variable is *exponentially distributed* with parameter $\beta = \lambda$.

Exercise

A certain widget has a mean time between failures of 24 hours, i.e., failures occur at a constant rate of one failure every 24 hours.
One evening, the widget was observed to be working at 10pm and then left unobserved for the night. The next morning at 6am, it was observed to have failed earlier. What is the probability that it was still working at 5 am that morning?

4.Stock Control Problem

Exercise

To determine the level to which the stock of some commodity should be allowed to fall before an order for a new batch is placed. The general procedure is to choose this level so as to give a specified probability of a stockout before the new batch arrives.
This probability P_n, known as the risk level, must hence be determined as a function of the reorder level n. It will, in fact, equal the probability that the number of demands during the lead time exceeds n.

4.Stock Control Problem

Exercise

We shall consider the two problems just described in the case where:

- The demand is random in the sense that the number of demands r during a fixed interval of time t has a Poisson distribution

$$\frac{e^{-\lambda t}(\lambda t)^r}{\Gamma(r+1)}$$

- The lead time t has a gamma type probability density function

$$\frac{\mu e^{-\mu t}(\mu t)^{k-1}}{\Gamma(k)}$$

With mean k/μ and variance k/μ^2.

End

Credit to Zhanpeng Zhou (TA of SP21)
Credit to Fan Zhang (TA of SU21)
Credit to Jiawen Fan (TA of SP21)
Credit to Zhenghao Gu (TA of SP20)