VE401 Probabilistic Methods in Eng.

RC#2 Discrete Distributions & Continuous Random Variables

> By TA: DONG Juechu, Mar. 2021

if you want to edit this note, you can find it here https://github.com/joydddd/VE401-2020SP-notes

Discrete Distributions

$X \sim \text{Bernoulli}(p)$

$$X : S \to \{0, 1\} \subset \mathbb{R}$$

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

Then X is said to be a **Bernoulli random variable** or follow a Bernoulli distribution with parameter p.

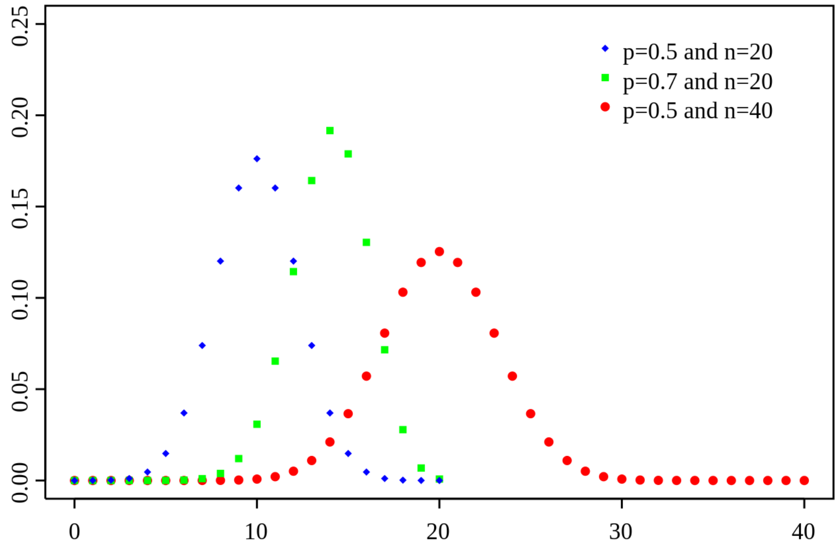
Binomial Distribution $X \sim B(n, p)$

Getting x success in n Bernoulli trails:

$$X : S \to \Omega = \{0, \ldots, n\} \subset \mathbb{R} \qquad (n \in \mathbb{N} \backslash \{0\})$$

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

Then X is said to be a **binomial random variable with parameters n and p.**



Geometric Distribution $X \sim \mathrm{Geom}(p)$.

Getting the first success on the xth Bernuolli trail:

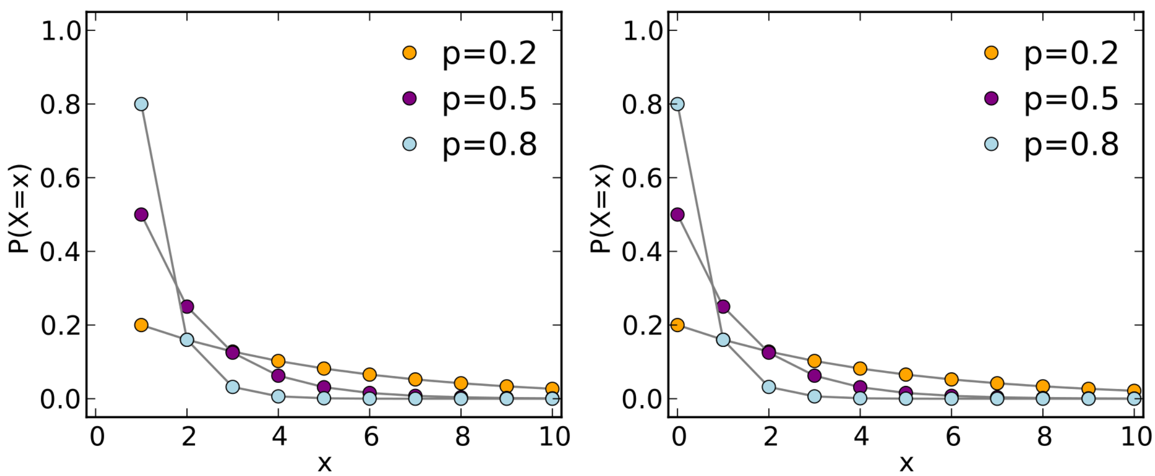
Let S be a sample space and

$$X : S \to \Omega = \mathbb{N} \backslash \{0\}$$

Let $0 < p < 1$ and define the density function $f_X : \mathbb{N} \backslash \{0\} \to \mathbb{R}$ given by

$$f_X(x) = (1-p)^{x-1} p.$$

We say that X is a **geometric random variable with parameter** p**.**



- moment generating function : $q = 1 - p$

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

- Expectation & Variance

$$E[X] = \frac{1}{p} \qquad Var[X] = \frac{q}{p^2}$$

Pascal Distribution

obtain the rth success at xth trail:

$$\begin{aligned} X : S \to \Omega &= \mathbb{N} \backslash \{0, 1, \ldots, r-1\} \\ &= \{r, r+1, r+2, \ldots\} \end{aligned} \qquad (r \in \mathbb{N} \backslash \{0\})$$

$$f_X : \Omega \to \mathbb{R}, \qquad f_X(x) = \binom{x-1}{r-1} p^r (1-p)^{x-r}, \quad 0 < p < 1$$

is said to follow a **Pascal distribution with parameters** p **and** r**.**

- <u>M.G.F.</u>

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

- <u>Expectation & Variance</u>

$$E[x] = \frac{r}{p} \qquad Var[X] = \frac{rq}{p^2}$$

- $$< Empty\ Math\ Block >$$

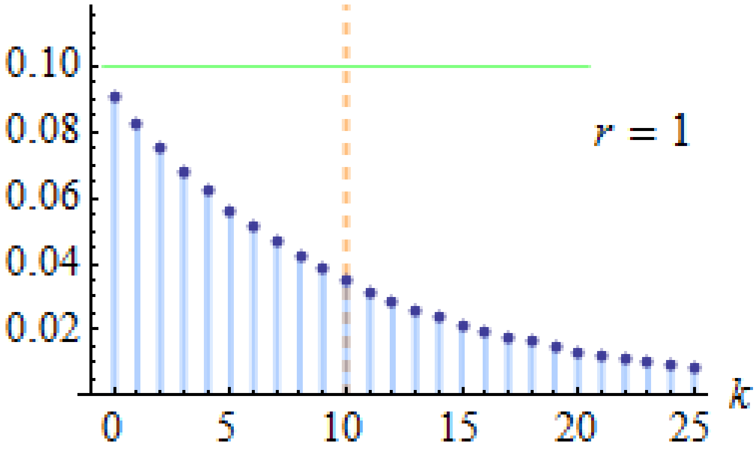
Negative binomial distribution

Experience x failures before the r success:

$$\begin{aligned} X : S \to \Omega &= \mathbb{N} \backslash \{0, 1, \ldots, r-1\} \\ &= \{r, r+1, r+2, \ldots\} \end{aligned}, \qquad r \in \mathbb{N} \backslash \{0\}$$

$$f_X : \Omega \to \mathbb{R}, \qquad f_X(x) = \binom{x-1}{r-1} p^r (1-p)^{x-r}, \quad 0 < p < 1$$

is said to follow a **Pascal distribution with parameters** p **and** r.



> Show that the sum of two independent and identical geometric random variables follows a Pascal distri-
> bution with r = 2.

Poison distribution

arrival in a continuous time interval:

Assumptions

1. Independence: If the intervals T1, T2 doesn't overlap (except perhaps at one point), then the numbers of arrivals in these interval are independent of each other.
2. Constant rate of arrivals

   Let $k = \lambda t, \lambda \text{ arriving rate}, t \text{ time interval}$

$$f_X(x) = \frac{k^x e^{-k}}{x!}$$

is said to have a **Poisson distribution with parameter k.**

- M.G.F.

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

- Expectation & Var

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

- cumulative distribution function

$$F(x) = P[X \leq x] = \sum_{y=0}^{\lfloor x \rfloor} \frac{e^{-k} k^y}{y!}$$

Poisson Approximation to the Binomial Distribution

Consider the density f of the binomial distribution,

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

Let k be fixed so that $np=k$ and set $p=k/n$. Replace p by k/n everywhere in $(*)$ and then let $n\to\infty$. Use Stirling's formula [1] to show that for every x, $f(x)\to(k^{x}/x!)\,e^{-k}$, the density of the Poisson distribution with parameter k.

[1] Stirling's formula states that

$$n!\sim\sqrt{2\pi n}\left(\frac{n}{e}\right)^{n}\qquad\text{as } n\to\infty$$

where $f(n)\sim g(n)$ as $n\to\infty$ means that $\lim_{n\to\infty}\frac{f(n)}{g(n)}=1$

Continuous Random Variable

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

(i) $f_X(x)\geq 0$ for all $x\in\mathbb{R}$ and

(ii) $\int_{-\infty}^{\infty}f_X(x)dx=1$.

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

$$P[a\leq X\leq b]=\int_a^b f_X(x)dx$$

The function f_X is called the **probability density function** (or just density) of the random variable X.

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

F_X is called the **cumulative distribution function**. We can obtain the density f_X from F_X :$f_X(x) = F'_X(x)$

$$
\begin{aligned}
&\mathrm{E}[X] := \int_{\mathbb{R}} x \cdot f_X(x) dx \\
&\mathrm{E}[\varphi \circ X] = \int_{-\infty}^{\infty} \varphi(x) \cdot f_X(x) dx \\
&\operatorname{Var}[X] := \mathrm{E}\left[(X - \mathrm{E}[X])^2\right] = \mathrm{E}\left[X^2\right] - \mathrm{E}[X]^2 \\
&m_X(t) = E\left[e^{tX}\right] = \int_{-\infty}^{\infty} e^{tx} f_X(x) dx
\end{aligned}
$$

- **median M_X:** defined by $P\left[X \leq M_X\right] = 0.5$. aka the time where half of the components will have failed
- **mean E[X].**
- **mode x_0,**: the location of the maximum of f_X (if there is a unique maximum location).aka. the time with the greatest failure density, i.e., the time around which failure is most likely. For the exponential distribution, $x_0 = 0$

Transformation of random variables

1.3.13. Theorem. Let X be a continuous random variable with density f_X. Let $Y = \varphi \circ X$, where $\varphi : \mathbb{R} \to \mathbb{R}$ is strictly monotonic and differentiable. The density for Y is then given by

$$
f_Y(y) = f_X\left(\varphi^{-1}(y)\right) \cdot \left|\frac{d\varphi^{-1}(y)}{dy}\right| \qquad \text{for } y \in \operatorname{ran} \varphi
$$

and

$$
f_Y(y) = 0 \qquad \text{for } y \notin \operatorname{ran} \varphi \,.
$$

Exponential Distribution

The probability of x arrivals in the time interval [0,t]

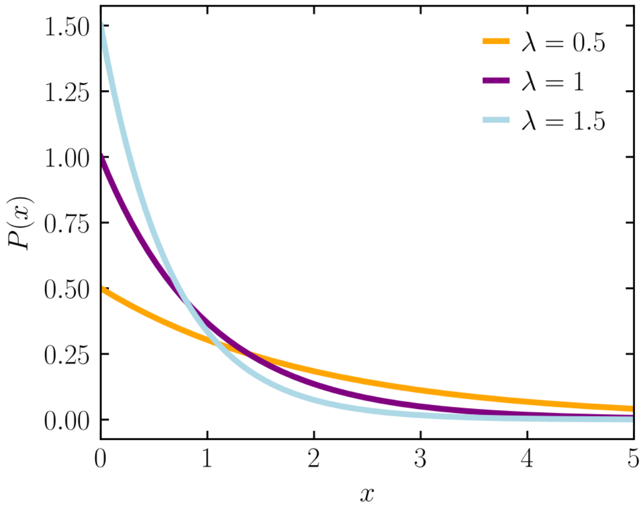
Time of the first arrival (continuous random variable) T

$F_T(t)$ the cumulative distribution on the density of T

$f_T(t) = \frac{d}{dt} F_T(t) = \lambda e^{-\lambda t}$ (exponential distribution with $\beta = \lambda$)

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

$$E[X] = \frac{1}{\beta} \qquad Var[X] = \frac{1}{\beta^1} \qquad m_X(t) = \frac{\beta}{\beta - t}$$



Gamma/Chi Square Distribution

The time of r arrivals

$F_{Tj}(t) = \lambda e^{-\lambda t} \frac{(\lambda t)^{j-1}}{(j-1)!}$

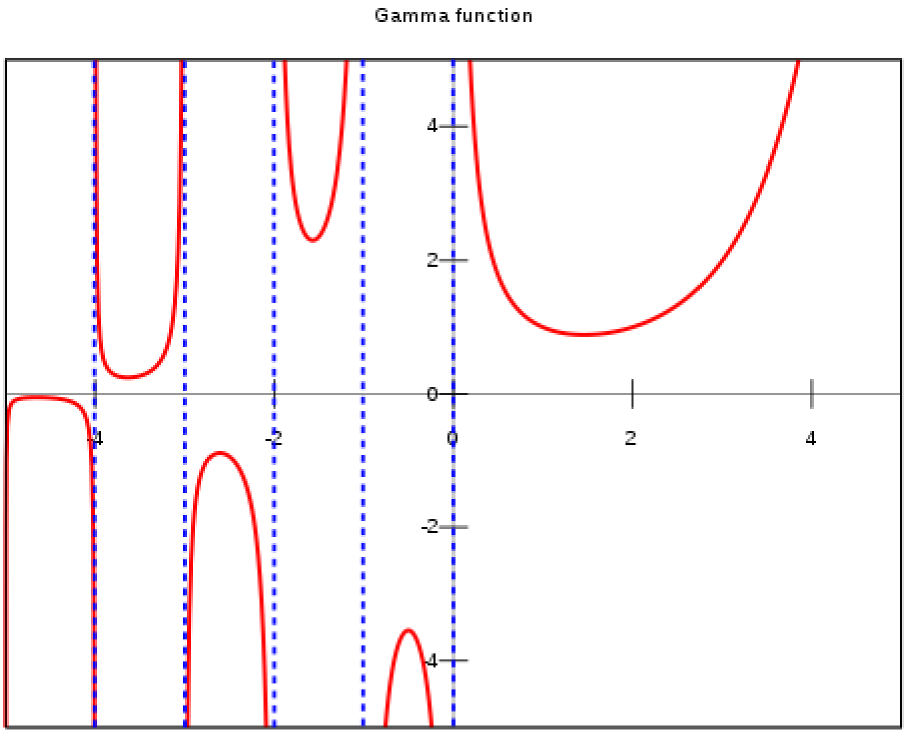
Let $\alpha, \beta \in \mathbb{R}, \alpha, \beta > 0$. A continuous random variable $(X, f_{\alpha,\beta})$ with density

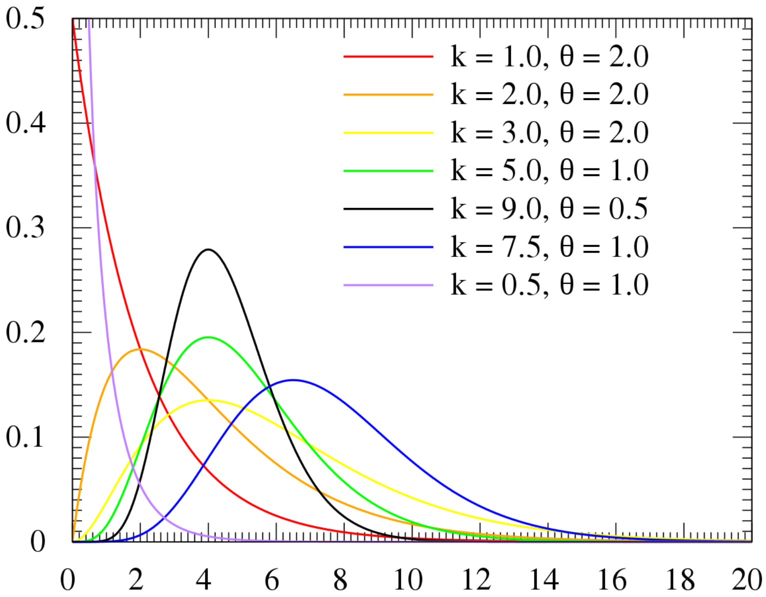
$$f_{\alpha,\beta}(x) = \begin{cases} \frac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha-1} e^{-x/\beta}, & x > 0 \\ 0, & x \le 0, \end{cases}$$

is said to follow an **gamma distribution with parameters α and β.**
Here

$$\Gamma(\alpha) = \int_0^\infty z^{\alpha-1} e^{-z} dz, \quad \alpha > 0$$

is the Euler gamma function. $(n! = \Gamma(n+1))$





- $E[X] = \alpha\beta \qquad VarX = \alpha\beta^2$
- M.G.F

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

Let $\gamma \in \mathbb{N}$. A continuous random variable $\left(\chi^2_\gamma, f_X\right)$ with density

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

is said to follow a **chi-squared distribution with γ degrees of freedom**. chi-squared distribution is simply that of a gamma random variable with $\beta = 2$ and $\alpha = \gamma/2$.

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

> ### Stock Control Problems
>
> https://www.jstor.org/stable/3007410?seq=7#metadata_info_tab_contents
>
> In order to ensure a high standard of serviceability without incurring unnecessary aircraft delay, an airline
> adopts the following replacement policy for its repairable aircraft components:
>
> 1. An unserviceable component is removed from an aircraft and sent for inspection to repair.
> 2. An immediate demand for a serviceable replacement is made to the stores. If available, the replacement
>   is at once fitted to the aircraft.
> 3. When the previously unserviceable component has been repaired to be serviceable, it is placed in the
>   stores.
>
> Suppose that the components are delivered serviceable to stores in the same order as they are removed
> from aircraft. We denote that
>
> - "Demand": a component for repair.
> - "Lead time" with parameter k: the time interval between the removal of the component and its
>   entry into stores, i.e., the time interval for the store to deliver an order of k specified components.
>
> What is the probability p_{rk} of exactly r demands during the lead time when the parameter value is
> k?