ECE4010J RC3

Shao Yujie

JOINT INSTITUTE
交大密西根学院

Continuous Random Variables, Exponential and Gamma distributions

- Continuous Random Variables：

The expectation of a continuous random variable is defined as follow:

$$E[X] := \int_{-\infty}^{+\infty} x f_X(x)dx$$

If the integral above converges absolutely.
The variance is defined as:

$$Var[X] := E[X^2] - E[X]^2$$

And for a given function φ, we have

$$E[\varphi(X)] = \int_{-\infty}^{+\infty} \varphi(x) f_X(x)dx$$

Continuous Random Variables, Exponential and Gamma distributions

- Exponential Distribution:

A continuous random variable with PDF $f_X(x) = \beta e^{-\beta x}(x \geq 0)$ is called an exponential distribution with parameter β.
Some important properties of exponential distribution:

- $E[X] = \frac{1}{\beta}$
- $Var[X] = \frac{1}{\beta^2}$
- m.g.f: $m_X(t) = (1 - \frac{t}{\beta})^{-1}$

- Memoryless Property

The exponential distribution has a special property called "memoryless property", i.e.,

$$P[X > s + t | X > s] = P[X > t]$$

And vice versa. It's the only continuous distribution that satisfy this property.
An LED's lifetime is likely to follow an exponential distribution. That is to say, when we have already used an LED for s hours and it's not broken, the expected lifetime is the same as a new LED. The LED won't "remember" that it has already used for s hours.
Human lifetime doesn't follow an exponential distribution. It's not memoryless.

Generalization: the time T_r needed for $r \in \mathbb{N} \setminus \{0\}$ arrivals to occur.

The cumulative distribution function is given by

$$\begin{aligned} F_{T_r}(t) &= P[T_r < t] \\ &= 1 - P[T_r > t] \\ &= 1 - P[\text{strictly less than } r \text{ arrivals before } t] \\ &= 1 - \sum_{n=0}^{r-1} \frac{(\lambda t)^n}{n!} e^{-\lambda t} \end{aligned}$$

for $t > 0$ and $F_{T_r}(t) = 0$ for $t < 0$.

```
In[2]:= PDF[ExponentialDistribution[λ], x]
```

$$\text{Out[2]=} \begin{cases} e^{-x\lambda}\lambda & x \geq 0 \\ 0 & \text{True} \end{cases}$$

Continuous Random Variables, Exponential and Gamma distributions

- Gamma Distribution:

$$f_{\alpha,\beta}(x) = \begin{cases} \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}, & x > 0, \\ 0, & x \leq 0, \end{cases} \qquad \Gamma(\alpha) = \int_0^\infty z^{\alpha-1} e^{-z}\, dz, \ \alpha > 0,$$

Has two parameter: α and β. It's PDF is tedious and you do not need to remember it. When $\alpha = 1$, this is the exponential distribution.
If X follows a Gamma distribution with parameter α and β,

- $E[X] = \frac{\alpha}{\beta}$
- $Var[X] = \frac{\alpha}{\beta^2}$
- m.g.f $m_X(t) = (1 - \frac{t}{\beta})^{-\alpha}$

The sum of α i.i.d random variable with parameter β is the Gamma distribution with parameter α and β. This can be seen from the m.g.f.

In[3]:= `PDF[GammaDistribution[α, β], x]`

Out[3]= $\begin{cases} \frac{e^{-\frac{x}{\beta}} x^{-1+\alpha} \beta^{-\alpha}}{\mathrm{Gamma}[\alpha]} & x > 0 \\ 0 & \mathrm{True} \end{cases}$

Here, /beta in mathematica is /beta ^(-1) in slide

- Chi-squared Distribution：

6.7. Definition. Let $\gamma \in \mathbb{N}$. A continuous random variable (χ^2_γ, f_X) with density

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

is said to follow a chi-squared distribution with γ ***degrees of freedom***.

The chi-squared distribution is simply a gamma distribution with $\beta = 1/2$ and $\alpha = \gamma/2$. It is worth noting that

$$\mathrm{E}[\chi^2_\gamma] = \gamma, \qquad \mathrm{Var}[\chi^2_\gamma] = 2\gamma.$$

This distribution plays an important role in statistics.

```
In[2]:= PDF[ChiSquareDistribution[ν], x]
```

$$\text{Out[2]=} \begin{cases} \dfrac{2^{-\nu/2} e^{-x/2} x^{-1+\frac{\nu}{2}}}{\mathrm{Gamma}\left[\frac{\nu}{2}\right]} & x > 0 \\ 0 & \mathrm{True} \end{cases}$$

The Normal Distribution

- The Normal Distribution:

7.1. Definition. Let $\mu \in \mathbb{R}$, $\sigma > 0$. A continuous random variable (X, f_X) with density

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-((x-\mu)/\sigma)^2/2}$$

is said to follow a normal distribution with parameters μ and σ.

7.3. Definition. A normally distributed random variable with parameters $\mu = 0$ and $\sigma = 1$ is called a ***standard normal*** random variable and denoted by Z.

7.4. Theorem. Let X be a normally distributed random variable with mean μ and standard deviation σ. Then

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

has standard normal distribution.

The Normal Distribution

- The Normal Distribution:

$$X_1 \sim N(\mu_1, \sigma_1^2), X_2 \sim N(\mu_2, \sigma_2^2) \Rightarrow X_1 + X_2 \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$$

- The Chebyshev Inequality

For a random variable X. If its expectation and variance exist, then

$$P[-m\sigma < X - \mu < m\sigma] \geq 1 - \frac{1}{m^2}$$

Since X can be any random variable, this estimation is very rough. However, it plays an important role in theoretical proof, like the proof of the weak law of large number.

The Normal Distribution

- Assuming that the total cholesterol in a population (in mmol /L) $X \sim N(4.2, 0.8)$, and that total cholesterol is considered high if it exceeds 5.6 mmol/L, find the probability that the total cholesterol is high.

- Assuming that the height of a bus door is designed in such a way that the chance of a man touching the top of the door is less than 0.01, and given that the height of a man (in centimeters) $X \sim N(175, 6)$, how is the height of the bus door designed?

- In statistics, the 68–95–99.7 rule, also known as the empirical rule, is a shorthand used to remember the percentage of values that lie within an interval estimate in a normal distribution: 68%, 95%, and 99.7% of the values lie within one, two, and three standard deviations of the mean, respectively.

The Normal Distribution

- Assuming that the total cholesterol in a population (in mmol /L) $X \sim N(4.2, 0.8)$, and that total cholesterol is considered high if it exceeds 5.6 mmol/L, find the probability that the total cholesterol is high.

```
In[7]:= CDF[NormalDistribution[0, 1], x]
```

$$\text{Out[7]=}\ \frac{1}{2}\,\text{Erfc}\left[-\frac{x}{\sqrt{2}}\right]$$

$$P\,(X > 5.6) = 1 - \frac{1}{2}\,\text{Erfc}\left[-\frac{\frac{5.6-4.2}{0.8}}{\sqrt{2}}\right]$$

$$\text{In[3]:=}\ 1 - \frac{1}{2}\,\text{Erfc}\left[-\frac{\frac{5.6-4.2}{0.8}}{\sqrt{2}}\right]$$

```
Out[3]= 0.0400592
```

The Normal Distribution

- Assuming that the height of a bus door is designed in such a way that the chance of a man touching the top of the door is less than 0.01, and given that the height of a man (in centimeters) X ~ N(175, 6), how is the height of the bus door designed?

let the height of boor be h

$P(X > h) \leq 0.01$ is the same as $P(X \leq h) \leq 0.99$

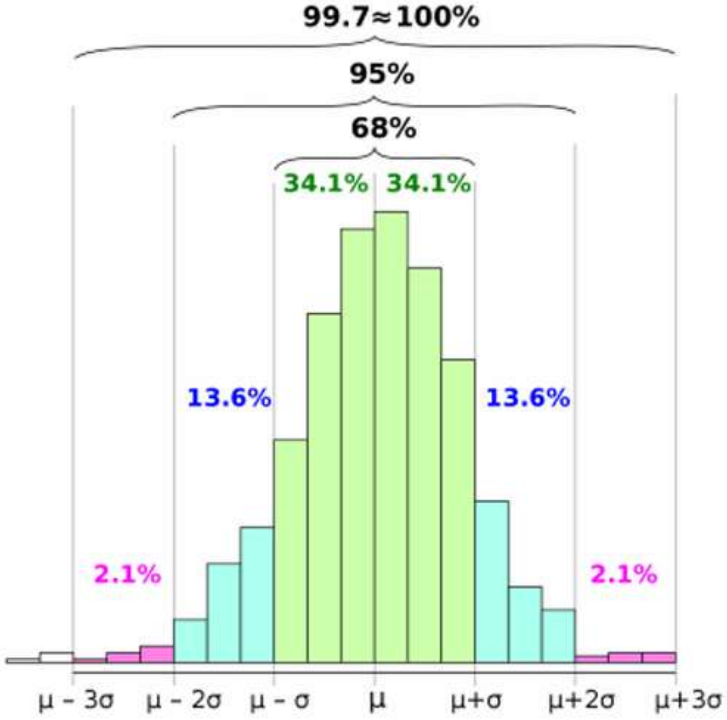
$$P(X \leq h) = \phi\left(\frac{h-\mu}{\sigma}\right) = \frac{1}{2}\,\mathrm{Erfc}\left[-\frac{\frac{h-175}{6}}{\sqrt{2}}\right] \geq 0.99$$

After looking up in the table, we can find that $\phi(2.33)=0.99$

So $h = 2.33*\sigma + \mu = 2.33*6 + 175 = 189$

The Normal Distribution

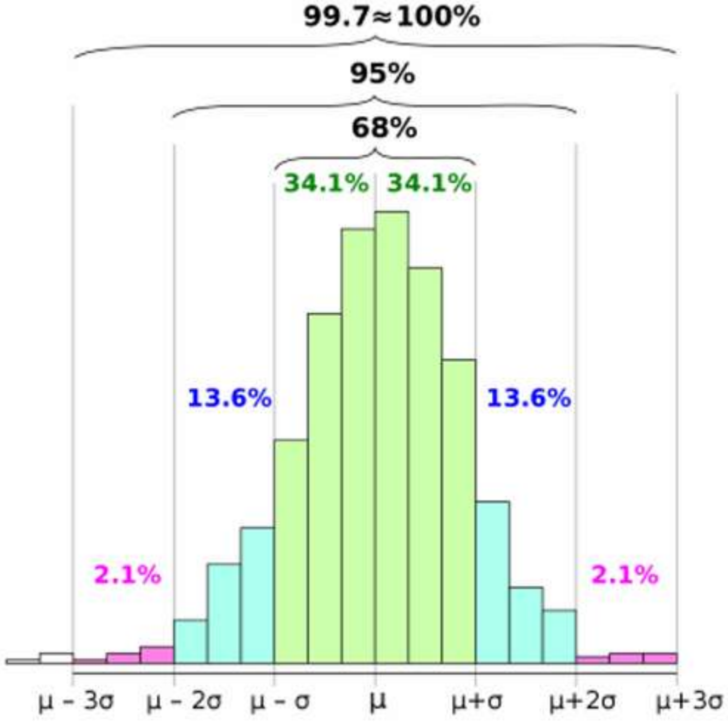
- In statistics, the 68–95–99.7 rule, also known as the empirical rule, is a shorthand used to remember the percentage of values that lie within an interval estimate in a normal distribution: 68%, 95%, and 99.7% of the values lie within one, two, and three standard deviations of the mean, respectively.



$$\Pr(\mu - 1\sigma \le X \le \mu + 1\sigma) \approx 68.27\%$$
$$\Pr(\mu - 2\sigma \le X \le \mu + 2\sigma) \approx 95.45\%$$
$$\Pr(\mu - 3\sigma \le X \le \mu + 3\sigma) \approx 99.73\%$$

The Normal Distribution

- In statistics, the 68–95–99.7 rule, also known as the empirical rule, is a shorthand used to remember the percentage of values that lie within an interval estimate in a normal distribution: 68%, 95%, and 99.7% of the values lie within one, two, and three standard deviations of the mean, respectively.



$$\Pr(\mu - 1\sigma \le X \le \mu + 1\sigma) \approx 68.27\%$$
$$\Pr(\mu - 2\sigma \le X \le \mu + 2\sigma) \approx 95.45\%$$
$$\Pr(\mu - 3\sigma \le X \le \mu + 3\sigma) \approx 99.73\%$$

In[1]:= `CDF[NormalDistribution[μ, σ], x]`

Out[1]= $\frac{1}{2}\,\mathrm{Erfc}\left[\frac{-x+\mu}{\sqrt{2}\,\sigma}\right]$

In[11]:= `1 - 2×N[CDF[NormalDistribution[μ, σ], μ - σ]]`

Out[11]= 0.682689

In[9]:= `1 - 2×N[CDF[NormalDistribution[μ, σ], μ - 2σ]]`

Out[9]= 0.9545

In[10]:= `1 - 2×N[CDF[NormalDistribution[μ, σ], μ - 3σ]]`

Out[10]= 0.9973

Distribution	P.D.F.	C.D.F.	M.G.F.	Expectation	Variance
Bernoulli	$f_X(x) = \begin{cases} 1-p, & x=0 \\ p, & x=1 \end{cases}$	–	–	p	$p(1-p)$
Binomial	$f_X(x) = \binom{n}{x} p^x (1-p)^{n-x}$	$F_X(x) = \sum_{y=0}^{\lfloor x \rfloor} \binom{n}{y} p^y (1-p)^{n-y}$	–	np	$np(1-p)$
Geometric	$f_X(x) = (1-p)^{x-1} p$	$F_X(x) = 1 - q^{\lfloor x \rfloor}$	$m_X(t) = \frac{pe^t}{1-qe^t}$	$\frac{1}{p}$	$\frac{1-p}{p^2}$
Pascal	$f_X(x) = \binom{x-1}{r-1} p^r (1-p)^{x-r}$	–	$m_X(t) = \frac{(pe^t)^r}{(1-qe^t)^r}$	$\frac{r}{p}$	$\frac{rq}{p^2}$
Negative Binomial	$f_X(x) = \binom{x+r-1}{r-1} p^r (1-p)^x$	–	–	–	–
Hypergeometric	$f_X(x) = \frac{\binom{r}{x}\binom{N-r}{n-x}}{\binom{N}{n}}$	–	–	$n\frac{r}{N}$	–
Poisson	$f_X(x) = \frac{k^x e^{-k}}{x!}$	$F_X(x) = \sum_{y=0}^{\lfloor x \rfloor} \frac{e^{-k} k^y}{y!}$	$m_X(t) = e^{k(e^t-1)}$	k	k
Exponential	$f_\beta(x) = \begin{cases} \beta e^{-\beta x}, & x > 0 \\ 0, & x \le 0 \end{cases}$	–	$m_X(t) = \frac{1}{1-\frac{t}{\beta}}$	$\frac{1}{\beta}$	$\frac{1}{\beta^2}$
Gamma	$f_{\alpha,\beta}(x) = \begin{cases} \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}, & x > 0 \\ 0, & x \le 0 \end{cases}$	–	$m_X(t) = \frac{1}{\left(1-\frac{t}{\beta}\right)^\alpha}$	$\frac{\alpha}{\beta}$	$\frac{\alpha}{\beta^2}$
Chi-Squared	$f_{\chi^2_n}(y) = \begin{cases} \frac{1}{2^{n/2}\Gamma(\frac{n}{2})} y^{n/2-1} e^{-y/2} & y > 0, \\ 0, & y \le 0. \end{cases}$	–	–	n	$2n$
Normal	$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$	–	$m_X(t) = e^{\mu t + \frac{\sigma^2 t^2}{2}}$	μ	σ^2
Standard Normal	$f_Z(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}$	–	$m_X(t) = e^{\frac{t^2}{2}}$	0	1
Student T	$f_{T_\gamma}(t) = \frac{\Gamma((\gamma+1)/2)}{\Gamma(\gamma/2)\sqrt{\pi\gamma}} \left(1+\frac{t^2}{\gamma}\right)^{-\frac{\gamma+1}{2}}$	–	–	–	–
F Ratio	$f_{\gamma_1,\gamma_2}(x) = \gamma_1^{\gamma_1/2} \gamma_2^{\gamma_2/2} \frac{\Gamma\left(\frac{\gamma_1+\gamma_2}{2}\right)}{\Gamma\left(\frac{\gamma_1}{2}\right)\Gamma\left(\frac{\gamma_2}{2}\right)} \frac{x^{\gamma_1/2-1}}{(\gamma_1 x+\gamma_2)^{(\gamma_1+\gamma_2)/2}}$	–	–	–	–

Approximation

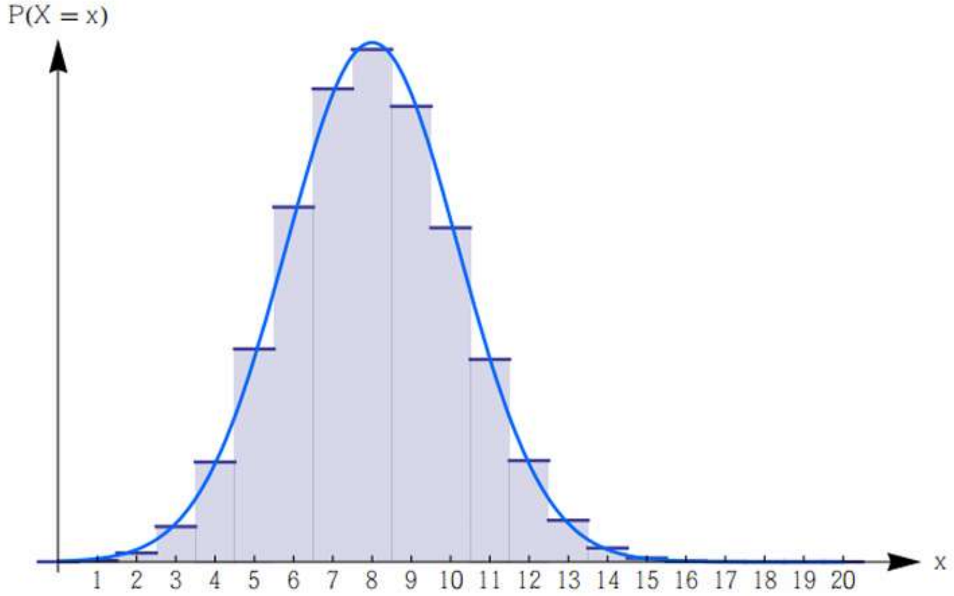
Usually, we can use the form of Chebyshev Inequality to do the approximation:

More generally, for $k \in \mathbb{N} \setminus \{0\}$,

$$P[|X| \geq c] \leq \frac{\mathrm{E}[|X|^k]}{c^k}. \qquad (7.1)$$

This is one version of ***Chebyshev's inequality***. The

We can also find the connection between Binomial Distribution and Normal Distribution:



To approximate in this way, we need to pay attention to the **half-unit correction**:

Hence, for $y = 0, \ldots, n$,

$$P[X \leq y] = \sum_{x=0}^{y} \binom{n}{x} p^x (1-p)^{n-x} \approx \Phi\left(\frac{y + 1/2 - np}{\sqrt{np(1-p)}}\right).$$

This additional term $1/2$ is known as the ***half-unit correction*** for the

This approximation is good if p is close to $1/2$ and $n > 10$. Otherwise, we require that

$$np > 5 \quad \text{if } p \leq 1/2 \quad \text{or} \quad n(1-p) > 5 \quad \text{if } p > 1/2.$$

BTW, the central limit theorem is also useful for approximation. It's an important theorem, which describes that under certain conditions (independent identically distributed with limited variance), the distribution of the sum of a large number of independent random variables will approach the normal distribution.

Let

$$Y_n = X_1 + \cdots + X_n.$$

Then for any $z \in \mathbb{R}$,

$$P\left[\frac{Y_n - \mathrm{E}[Y_n]}{\sqrt{\mathrm{Var}[Y_n]}} \leq z\right] \xrightarrow{n \to \infty} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-x^2/2}\, dx.$$

$$Z_n = \frac{S_n - n\mu}{\sigma\sqrt{n}}$$

Thank You !!!

From The Elder Scrolls V: Skyrim Special Edition