Parametric continuous distributions

Li Li

Department of Mathematics and Statistics
Continuous uniform distribution

A continuous uniform random variable $X$ has probability density function (PDF)

$$f(x) = \frac{1}{b - a}, \quad a \leq x \leq b,$$

and 0 elsewhere. Here $a$ and $b$ are two real numbers. We refer the distribution as uniform $(a, b)$.

- For example, a continuous uniform distribution over $[0, 1]$ (often referred as uniform $(0, 1)$) has density $f(x) = 1$ for $0 < x < 1$ and 0 elsewhere.
- Interesting fact: given any continuous random variable $Y$ and its cumulative distribution function $F(y)$, a CDF transformation of $Y$, i.e $F(Y)$ has a uniform $(0, 1)$ distribution.
- For example, if $Y$ has a density $f(y) = \lambda e^{-\lambda y}$ for $y \geq 0$, then its CDF is $F(y) = 1 - e^{-\lambda y}$ for $y \geq 0$. We have $1 - e^{\lambda Y} \sim \text{Uniform}(0, 1)$. 
\[
\begin{align*}
\frac{1}{b-a} & \quad \text{Uniform distribution on } [a, b] \\
\int_{c}^{d} \frac{1}{b-a} \, dx & = \int_{c}^{d} \frac{d-c}{b-a} \\
\P(a-x < x^* = \frac{x^*-a}{b-a} \\
X & \sim \text{Uniform } (0, 1) \\
F(.): \text{ CDF function} \\
Y: F^{-1}(X) \quad \text{Y has distribution } F(.)
\end{align*}
\]
Where is continuous uniform distribution used?

• Simulations are used to model complicated processes, estimate distributions of estimators (using methods such as bootstrap), and have dramatically increased the use of an entire field of statistics.

• In simulations, we often generate random numbers from a desired distribution.

• Uniform \((0, 1)\) is the where random number generation start.

• To generate a random variable that has CDF \(F(y) = 1 - e^{-\lambda y}\) for \(y \geq 0\), we can use the following steps
  
  (a) generate a random number \(u\) from Uniform \((0, 1)\).
  
  (b) transform the random number \(u\) by the inverse of the CDF for the density we desire, i.e. \(-\log(1 - u)/\lambda\).
CDF of Uniform \((a, b)\)

\[
F(x) = \begin{cases} 
0 & x \leq a \\
\frac{x-a}{b-a} & a < x < b \\
1 & x = b
\end{cases}
\]

\[
F(b) = \frac{b-a}{b-a} = 1
\]
Mean and variance of uniform \((a, b)\):

\[ b^2 - a^2 = (b-a)(b^2 + ba + a^2) \]

\[ \mu = \frac{a+b}{2} \]

\[ \sigma^2 = \frac{(b-a)^2}{12} \]

\[ \mu = \int_{a}^{b} x \cdot \frac{1}{b-a} \, dx \]

\[ = \left. \frac{x^2}{2(b-a)} \right|_{a}^{b} \]

\[ = \frac{b^2-a^2}{2(b-a)} = \frac{b+a}{2} \]

\[ \sigma^2 = E(X^2) - \mu^2 \]

\[ = \int_{a}^{b} \frac{x^2}{b-a} \, dx - \left( \frac{b+a}{2} \right)^2 \]

\[ = \left. \frac{x^3}{3(b-a)} \right|_{a}^{b} - \frac{(b+a)^2}{4} \]

\[ = \frac{b^2+ba+a^2}{3} - \frac{(b+a)^2}{4} \]
\[ f(x) = \begin{cases} \ f_1(x) & -1 < x < 0 \\ \ f_2(x) & 0 \leq x < 1 \end{cases} \]

\[ E(x) = \int_{-\infty}^{+\infty} x \cdot f(x) \, dx = \int_{-1}^{0} x \cdot f_1(x) \, dx + \int_{0}^{1} x \cdot f_2(x) \, dx \]
Normal distribution

1) Characteristics

2) \( P( a < X < b ) \)

3) Percentiles

4) Normal approximations to Binomial distribution and Poisson distribution
Normal distributions

Given parameters $\mu$ and $\sigma$, PDF of the Normal distribution is

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

Figure: Normal distribution density functions. Density function is denoted as $\phi_{\mu,\sigma^2}(x)$.

- $\mu$–location parameter or mean of the distribution and $\sigma$–scale parameter or standard deviation of the distribution.
• A Normal distribution with \( \mu = 0 \) and \( \sigma = 1 \) is referred as “standard Normal distribution”.
• Normal distributions are also called ”Gaussian distributions”.
• Normal distribution is sometimes informally called the “bell curve”.
• A random variable with a Gaussian distribution is said to be normally distributed, denoted by \( X \sim \mathcal{N}(\mu, \sigma^2) \). For example, \( X \sim \mathcal{N}(3, 2^2) \).
Important facts about Normal distributions

- All Normal curves have the same overall shape: symmetric, single-peaked, bell-shaped.
- Any specific Normal curve is completely described by giving its mean $\mu$ and its standard deviation $\sigma$.
- The mean is located at the center of the symmetric curve and is the same as the median. Changing $\mu$ without changing $\sigma$ moves the Normal curve along the horizontal axis without changing its spread.
- The standard deviation $\sigma$ controls the spread of a Normal curve. Curves with larger standard deviations are more spread out or wider.
Important facts about Normal distributions

• The average of many independent processes (such as measurement errors) often have distributions that are nearly normal.

• If $X_1, X_2, \ldots, X_n$ are independent Bernoulli random variables with the same success rate $p$, $\bar{X}$ (average of $X_1, \ldots, X_n$) follows a Normal distribution $N(p, \frac{p(1-p)}{n})$ approximately.

• If $X_1, X_2, \ldots, X_n$ are independent Poisson random variables with the same rate of event $\lambda$ and time interval $T$, $\bar{X}$ follows a Normal distribution $N(\lambda T, \frac{\lambda T}{n})$ approximately.
For a Normal distribution with mean $\mu$ and standard deviation $\sigma$:

- Approximately 68% of the observations fall within $\sigma$ of the mean $\mu$.
- Approximately 95% of the observations fall within $2\sigma$ of the mean $\mu$.
- Approximately 99.7% of the observations fall within $3\sigma$ of the mean $\mu$.  

The empirical rule

**Figure:** The 68–95–99.7 rule.

- $P(X > \mu + 3\sigma) = \frac{0.0027}{2} \approx 0.00135$
- $P(X > \mu + \sigma) = \frac{(1 - 0.6827) - 1}{2} \approx 0.05 - \frac{0.6827}{2}$
- $P(X > \mu + 1.5\sigma)$

- $P(\mu - \sigma < X < \mu + \sigma) \approx 68.27\%$
- $P(\mu - 2\sigma < X < \mu + 2\sigma) \approx 95.45\%$
- $P(\mu - 3\sigma < X < \mu + 3\sigma) \approx 99.73\%$

- $P(X < \mu - 3\sigma) = 1 - 99.73\% = 0.0027$
How do we find $P(a < X < b)$

For a Normal distribution with mean $\mu$ and variance $\sigma^2$ and $X \sim N(\mu, \sigma^2)$, probability

$$P(a < X < b) = \int_a^b \frac{1}{\sqrt{2\pi \sigma^2}} e^{-(x-\mu)^2/2\sigma^2} \, dx$$

can be calculated by $\text{pnorm}((b - \mu)/\sigma) - \text{pnorm}((a - \mu)/\sigma)$ in R.

$\text{pnorm}(x)$: returns $F(x)$ where $F(\cdot)$ is the CDF of Standard Normal distribution.

$$P(a < X < b) = F\left(\frac{b-\mu}{\sigma}\right) - F\left(\frac{a-\mu}{\sigma}\right)$$
How do we find the $p$-th percentile?

$p$-th percentile of a continuous distribution with density $f(x)$:

Solve $F(c) = \int_{-\infty}^{c} f(x) \, dx = \frac{p}{100}, \quad 0 \leq p \leq 100$

The $p$-th percentile $c$ where

$$P(X < c) = \int_{\infty}^{c} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \, dx = \frac{p}{100}$$

can be calculated by $\text{qnorm}(\frac{p}{100}) \ast \sigma + \mu$. 

\[13\]
The cumulative distribution function (CDF) of a Normal distribution function is denoted as

\[
\Phi_{\mu, \sigma^2}(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi} \sigma^2} e^{-\frac{(u-\mu)^2}{2 \sigma^2}} \, du.
\]

Figure: Normal distribution CDF functions. CDF functions are denoted as \(\Phi_{\mu, \sigma^2}(x)\).
Standard Normal distribution and z-score

Standard Normal distribution:

- If a variable \(X\) has any Normal distribution \(N(\mu, \sigma^2)\), then the standardized variable \(Z = \frac{X - \mu}{\sigma}\) has the standard Normal distribution \(N(0, 1)\). For example, if \(X\) follows a Normal distribution with mean 3 and variance 4, i.e. \(X \sim N(3, 2^2)\), then \(Z = \frac{X - \mu}{\sigma} \sim N(0, 1)\).

- For a real number \(a\), the standardized value of \(a\) is \(z = \frac{a - \mu}{\sigma}\) and is called the z-score of \(a\).
\[
F_z(z) = P(Z \leq z)
\]

\[
= P\left( \frac{X - \mu}{\sigma} \leq z \right)
\]

\[
= P\left( X \leq \sigma z + \mu \right)
\]

\[
= \int_{-\infty}^{\sigma z + \mu} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \, dx
\]

Show: Let \( y = \frac{x-\mu}{\sigma} \)

\[
= \int_{-\infty}^{\sigma z + \mu} \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} \, dy
\]

\[
= F_{0.1}(z)
\]
Use of $z$-score

- When the $z$-score of an observation has an absolute value greater than 3, this observation can be viewed roughly as an outlier or unusual.

- $z$-score can be used to compare two observations from two populations that have different Normal distributions.
Example

Consider for two high school senior students,

- student A scored 670 on the Mathematics part of the SAT. The distribution of SAT Math scores in 2010 was Normal with mean 516 and standard deviation 116.
- student B took the ACT and scored 46 on the Mathematics portion. ACT Math scores for 2010 were Normally distributed with mean 21.0 and standard deviation of 5.3.

(a) Find the $z$-scores for both students.

(b) Assuming that both tests measure the same kind of ability, who had a higher score? Are any of these two test scores outlying?
a) \[ Z_A = \frac{670 - 516}{116} = 1.3276 \]

\[ Z_B = \frac{46 - 21}{5.3} = 4.7170 \]
Normal table

Goal: \( X \sim N(\mu, \sigma^2) \); \( P(a < X < b) \) from the Normal table.

- Only used in test situation these days.
- It is a one to one mapping of \( z \) to \( \Phi_{0,1}(z) \) (Standard Normal CDF) for \( z \) goes from \(-3.99\) to \(3.99\).
- Given \( z \), use the table we can find \( \Phi_{0,1}(z) \). Given \( p \) such that \( \Phi_{0,1}(z) = p \), we can also find \( z \). This gives us the \( p \times 100\)-th percentile of the Standard Normal distribution.
**Figure:** One to one mapping of $z$ to $\Phi_{0.1}(z)$ for $z$ from $-3.99$ to $0$. 

<table>
<thead>
<tr>
<th>$z$</th>
<th>$\Phi_{0.1}(z)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>-3.91</td>
<td>$0.0000$</td>
</tr>
<tr>
<td>-3.92</td>
<td>$0.0001$</td>
</tr>
<tr>
<td>-0.00013</td>
<td>$z$</td>
</tr>
</tbody>
</table>

$\Phi_{0.1}(z) = 0.00013$
Normal table

<table>
<thead>
<tr>
<th>Z</th>
<th>.00</th>
<th>.01</th>
<th>.02</th>
<th>.03</th>
<th>.04</th>
<th>.05</th>
<th>.06</th>
<th>.07</th>
<th>.08</th>
<th>.09</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0</td>
<td>50.000</td>
<td>50.399</td>
<td>50.798</td>
<td>51.197</td>
<td>51.595</td>
<td>51.994</td>
<td>52.392</td>
<td>52.790</td>
<td>53.188</td>
<td>53.586</td>
</tr>
<tr>
<td>0.1</td>
<td>53.586</td>
<td>53.880</td>
<td>54.176</td>
<td>54.472</td>
<td>54.767</td>
<td>55.062</td>
<td>55.356</td>
<td>55.650</td>
<td>55.944</td>
<td>56.238</td>
</tr>
<tr>
<td>0.2</td>
<td>57.926</td>
<td>58.317</td>
<td>58.706</td>
<td>59.095</td>
<td>59.483</td>
<td>59.871</td>
<td>60.257</td>
<td>60.642</td>
<td>61.026</td>
<td>61.409</td>
</tr>
<tr>
<td>0.3</td>
<td>61.791</td>
<td>62.172</td>
<td>62.552</td>
<td>62.930</td>
<td>63.307</td>
<td>63.683</td>
<td>64.058</td>
<td>64.431</td>
<td>64.803</td>
<td>65.173</td>
</tr>
<tr>
<td>0.4</td>
<td>65.542</td>
<td>65.910</td>
<td>66.276</td>
<td>66.640</td>
<td>67.003</td>
<td>67.364</td>
<td>67.724</td>
<td>68.082</td>
<td>68.439</td>
<td>68.793</td>
</tr>
<tr>
<td>0.5</td>
<td>69.146</td>
<td>69.497</td>
<td>69.847</td>
<td>70.194</td>
<td>70.540</td>
<td>70.884</td>
<td>71.226</td>
<td>71.566</td>
<td>71.904</td>
<td>72.240</td>
</tr>
<tr>
<td>0.6</td>
<td>72.757</td>
<td>73.097</td>
<td>73.437</td>
<td>73.775</td>
<td>74.113</td>
<td>74.451</td>
<td>74.789</td>
<td>75.127</td>
<td>75.465</td>
<td>75.802</td>
</tr>
<tr>
<td>0.7</td>
<td>75.804</td>
<td>76.135</td>
<td>76.464</td>
<td>76.793</td>
<td>77.122</td>
<td>77.450</td>
<td>77.778</td>
<td>78.106</td>
<td>78.434</td>
<td>78.762</td>
</tr>
<tr>
<td>0.8</td>
<td>78.814</td>
<td>79.103</td>
<td>79.389</td>
<td>79.673</td>
<td>79.955</td>
<td>80.234</td>
<td>80.511</td>
<td>80.785</td>
<td>81.057</td>
<td>81.327</td>
</tr>
<tr>
<td>0.9</td>
<td>81.504</td>
<td>81.595</td>
<td>81.812</td>
<td>82.019</td>
<td>82.216</td>
<td>82.404</td>
<td>82.592</td>
<td>82.779</td>
<td>82.966</td>
<td>83.153</td>
</tr>
</tbody>
</table>

Figure: One to one mapping of z to \( \Phi_{0.1}(z) \) for z from 0 to 3.99.
What is the probability for a standard Normal variable $Z$ take values less than 1.47?

- Locate 1.4 in the left-hand column of the Normal table
- Then locate the remaining digit seven as .07 in the top row.
- The entry opposite 1.4 and under .07 is 0.9292. This is the cumulative proportion we seek.

What is the 92.9-th percentile of a Standard Normal distribution?—It is 1.47.
For a Normal random variable follows a distribution that has a mean $\mu$ and variance $\sigma^2$, how to find the probability for the random variable to fall within an interval $[a, b]$?

- Note that $P(a < X < b) = P\left(\frac{a-\mu}{\sigma} < \frac{X-\mu}{\sigma} < \frac{b-\mu}{\sigma}\right)$ and $\frac{X-\mu}{\sigma}$ follows a standard Normal distribution.
- Hence $P(a < X < b) = \Phi_{0,1}\left(\frac{b-\mu}{\sigma}\right) - \Phi_{0,1}\left(\frac{a-\mu}{\sigma}\right)$.
- Calculate $\frac{b-\mu}{\sigma}$ and $\frac{a-\mu}{\sigma}$. Then use the Normal table to find the corresponding probabilities $\Phi_{0,1}\left(\frac{b-\mu}{\sigma}\right)$ and $\Phi_{0,1}\left(\frac{a-\mu}{\sigma}\right)$ respectively.
Suppose $X \sim N(2, 2^2)$, show that $Y \sim N(3, 1^2)$.

$P(1 < X < 3) = \Phi(0.5) - \Phi(-0.5) = 0.383.$

\[
P(1 < x < 3) = P\left(\frac{1-2}{2} < \frac{x-2}{2} < \frac{3-2}{2}\right)
= P\left(-0.5 < z < 0.5\right) \quad z \sim N(0, 1)
\]

\[
= \Phi(0.5) - \Phi(-0.5)
\]

\[
P(a < z < b) = P(z < b) - P(z < a)
\]
Finding percentiles using the Normal table

For a Normal random variable follows a distribution that has a mean $\mu$ and standard deviation $\sigma$, how to find the the $p$-th percentile of the distribution?

(a) Note that our goal is to find $c$ such that $P(X < c) = p/100$. Since $P(X < c) = \Phi_{0,1}(\frac{c-\mu}{\sigma})$, we are solving $c$ from

$$\Phi_{0,1}(\frac{c-\mu}{\sigma}) = \frac{p}{100}.$$ 

(b) Use the table to find the $p$-th percentile of the standard Normal distribution. Denoted it by $z_p$.

(c) Then $c = z_p \sigma + \mu$. 


Suppose $X \sim N(2, 2^2)$, show that the 92.9-th percentile of the distribution is $1.47 \times 2 + 2 = 4.94$.

$\Phi(1.47) = 92.9$

Goal: find $p$-th percentile of distribution $N(\mu, \sigma^2)$

i.e. find $c$ such that

$$P(X \leq c) = \frac{p}{100}$$

where $X \sim N(\mu, \sigma^2)$

$$\Phi_{\frac{c-\mu}{\sigma}} = \frac{p}{100}$$

$$\frac{c-\mu}{\sigma} = z_p$$

$$c = z_p \cdot \sigma + \mu$$
How to tell whether observations from a population follows a Normal distribution? (Chapter 7)

- Normal probability plot or QQ plot.
- Shapiro-Wilk test.
Normal approximation to the Binomial distribution: If $X$ is a binomial random variable with parameters $n$ and $p$,

$$Z = \frac{X - np}{\sqrt{np(1 - p)}}$$

follows a standard Normal distribution approximately. The approximation is close if $np > 5$ and $n(1 - p) > 5$.

[https://newonlinecourses.science.psu.edu/stat414/node/179/](https://newonlinecourses.science.psu.edu/stat414/node/179/)
Normal approximation to Binomial distributions

- To approximate a binomial probability with a normal distribution, a continuity correction is applied as follows:

\[
P(X \leq x) = P(X \leq x + 0.5) \approx P\left(Z \leq \frac{x + 0.5 - np}{\sqrt{np(1-p)}}\right)
\]

and

\[
P(X \geq x) = P(X \geq x - 0.5) \approx P\left(Z \geq \frac{x - 0.5 - np}{\sqrt{np(1-p)}}\right)
\]

- For example, if \(n = 20\) and \(p = 0.3\), then
  \[P(X \leq 7) \approx \Phi_{0,1}\left(\frac{7 + 0.5 - 6}{\sqrt{4.2}}\right) = 0.758,\]
  \[P(X \geq 7) \approx 1 - \Phi_{0,1}\left(\frac{7 - 0.5 - 6}{\sqrt{4.2}}\right) = 1 - 0.596 = 0.404.\]

- Note that \(P(X < x) = P(X \leq x - 1)\) and \(P(X > x) = P(X \geq x + 1)\).
Example

Assume that in a digital communication channel, the number of bits received in error can be modeled by a Binomial random variable, and assume that the probability that a bit is received in error is $1 \times 10^{-5}$. If 16 million bits are transmitted, what is the probability that 150 or fewer errors occur?
Normal approximation to the Poisson distribution: If $X$ is a Poisson random variable with $E(X) = \lambda T$ and $V(X) = \lambda T$,

$$P(X \leq x) = P(X \leq x + 0.5) = P\left(Z \leq \frac{x + 0.5 - \lambda T}{\sqrt{\lambda T}}\right)$$

and

$$P(X \geq x) = P(X \geq x - 0.5) = P\left(Z \geq \frac{x - 0.5 - \lambda T}{\sqrt{\lambda T}}\right)$$

The approximation is generally good for $\lambda T > 5$. 
Chi-square distributions

If $Z$ follows a standard Normal distribution then:

$$V = Z^2 \sim \chi_1^2,$$

where $\chi_1^2$ is called a chi-square distribution with 1 degree of freedom which has density

$$f(v) = \frac{1}{\Gamma(1/2)2^{0.5}} v^{0.5-1} e^{-v/2}, \ v \geq 0.$$
Chi-square distributions

More generally, if $Z_1, Z_2, \ldots, Z_k$ are independent (one does not affect the distribution of another), standard Normal random variables, then

$$V = \sum_{i=1}^{k} Z_i^2 \sim \chi_k^2$$

which denotes a chi-squared distribution with $k$ degrees of freedom. For example, $V = Z_1^2 + Z_2^2 \sim \chi_2^2$. The density of the distribution is

$$f(v) = \frac{1}{\Gamma(k/2)2^{k/2}} v^{k/2-1} e^{-v/2}, v \geq 0.$$  

Chi-squared distributions are used primarily in hypothesis testing.
Distributions of random sample mean and random sample variance

Suppose $X_1, \ldots, X_n$ are independent Normal random variables, which all follow distribution $N(\mu, \sigma^2)$, which means they are identical.

- Denote $\bar{X} = \frac{1}{n} \sum_{i=1}^{n} X_i$ as the random sample mean.
- Denote $S^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X})^2$ as the random sample variance.

Then the results are true regardless what values of $\mu$ and $\sigma^2$ are:

- $\bar{X} \sim N(\mu, \sigma^2/n)$
- $(n - 1)S^2/\sigma^2 \sim \chi^2_{n-1}$
Example

Suppose $X_1, \ldots, X_{20}$ are independent Normal random variables, which all follow distribution $N(2, 3^2)$, which means they are identical.

- Denote $\bar{X} = \frac{1}{20} \sum_{i=1}^{20} X_i$ as the random sample mean.
- Denote $S^2 = \frac{1}{19} \sum_{i=1}^{20} (X_i - \bar{X})^2$ as the random sample variance.

- $\bar{X} \sim N(2, 9/20)$
- $19S^2/9 \sim \chi^2_{19}$
Simple random sample

If $X_1, \ldots, X_n$ is called a simple random sample if

- $X_1, X_2, \ldots, X_n$ are independent random variables.
- $X_1, X_2, \ldots, X_n$ follow the same distribution, i.e. they are identical.
Central limit theorem (CLT)

If \( X_1, \ldots, X_n \) is a random sample of size \( n \) taken from a population or a distribution (not necessarily Normal distribution) with mean \( \mu \) and variance \( \sigma^2 \) and if \( \bar{X} \) is the sample mean, then

\[
\bar{X} \sim N(\mu, \sigma^2/n)
\]

for large \( n \). For example,

- If \( X_1, X_2, \ldots, X_{10} \) are independent random variables following an uniform distribution \((0, 1)\), then \( \bar{X} \) follows a Normal distribution \( N(0.5, 1/12/10) \), i.e. \( N(0.5, 0.0083) \).

- If \( X_1, X_2, \ldots, X_n \) are independent Bernoulli random variables with the same success rate 0.4, \( \bar{X} \) (average of \( X_1, \ldots, X_n \)) follows a Normal distribution \( N(0.4, \frac{0.24}{n}) \) approximately.
Animation of CLT

https://www.youtube.com/watch?v=PujollyCl_A