Lecture 7: Continuous Random Variable

Donglei Du
(ddu@unb.edu)

Faculty of Business Administration, University of New Brunswick, NB Canada Fredericton
E3B 9Y2

Table of contents

Layout

1. Continuous Random Variable
   - Probability Density Function (pdf)
   - Probability of any set of real numbers
2. Normal Random Variable
   - Standard Normal Random Variable
   - General Normal Random Variable
3. Relationship between $Z \sim N(0,1)$ and $X \sim N(\mu, \sigma^2)$
4. Calculations with Standard Normal Random Variable via the Normal Table
   - Given z-value, calculate probability
   - Given probability, calculate z-value
5. Calculations with General Normal Random Variable via the Normal Table
   - Given x-value, calculate probability
   - Given probability, calculate x-value

Continuous Random Variable

- A continuous random variable is any random variable whose set of all the possible values is uncountable.

Probability Density Function (pdf)

- A probability density function (pdf) for any continuous random variable is a function $f(x)$ that satisfies the following two properties:
  (i) $f(x)$ is nonnegative; namely,
  $$f(x) \geq 0$$
  (ii) The total area under the curve defined by $f(x)$ is 1; namely
  $$\int_{-\infty}^{\infty} f(x)dx = 1$$

Probability of any set of real numbers

- Given a continuous random variable X with its probability density function $f(x)$, for any set B of real numbers, the probability of B is given by

$$P(X \in B) = \int_B f(x)dx$$

- For instance, if $B = [a, b]$, then the probability of B is given by

$$P(a \leq X \leq b) = \int_a^b f(x)dx$$

- Geometrically, the probability of B is the area under the curve $f(x)$.

Example

- Consider the continuous random variable X with its probability density function $f(x)$ defined below

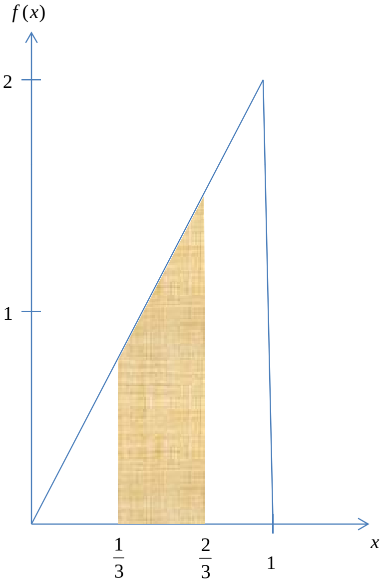
$$f(x) = \begin{cases} 2x, & 0 \leq x \leq 1 \\ 0, & x > 1 \end{cases}$$

- For instance, the probability of $[1/3, 2/3]$ is given by

$$P(1/3 \leq X \leq 2/3) = \int_{1/3}^{2/3} 2x dx = (2/3)^2 - (1/3)^2 = 1/3.$$

- Geometrically, the probability of $[1/3, 2/3]$ is the area under the curve $f(x)$ between $[1/3, 2/3]$.

Example



Note

- When dealing with a continuous random variable, we assume that the probability that the variable will take on any particular value is 0! Instead, probabilities are assigned to intervals of values!
- Therefore, give a continuous random variable X, then for any constant a:

$$P(X = a) = 0$$

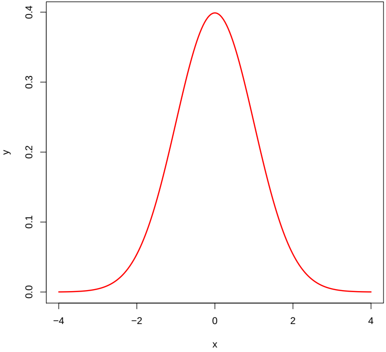
Layout

1. Continuous Random Variable
   - Probability Density Function (pdf)
   - Probability of any set of real numbers
2. Normal Random Variable
   - Standard Normal Random Variable
   - General Normal Random Variable
3. Relationship between $Z \sim N(0,1)$ and $X \sim N(\mu, \sigma^2)$
4. Calculations with Standard Normal Random Variable via the Normal Table
   - Given z-value, calculate probability
   - Given probability, calculate z-value
5. Calculations with General Normal Random Variable via the Normal Table
   - Given x-value, calculate probability
   - Given probability, calculate x-value

Standard Normal Random Variable

- The standard normal random variable Z has the following probability density function:

$$\phi(z) = \frac{1}{\sqrt{2\pi}}\, e^{-\frac{1}{2}z^2}.$$



Standard normal curve: Plot R code

```
> x<-seq(-4,4,length=200)
>y<-dnorm(x,mean=0,sd=1)
> plot(x,y,type="l",lwd=2,col="red")
```

Properties of Standard Normal Random Variable

- The pdf is symmetric around its mean $x = 0$, which is at the same time the mode, the median of the distribution.
- It is unimodal.
- It has inflection points at $+1$ and -1.
- Z has zero mean and unit variance; namely

$$\begin{aligned} \mathbb{E}[Z] &= 0 \\ \mathbb{V}[Z] &= 1 \end{aligned}$$

General Normal Random Variable

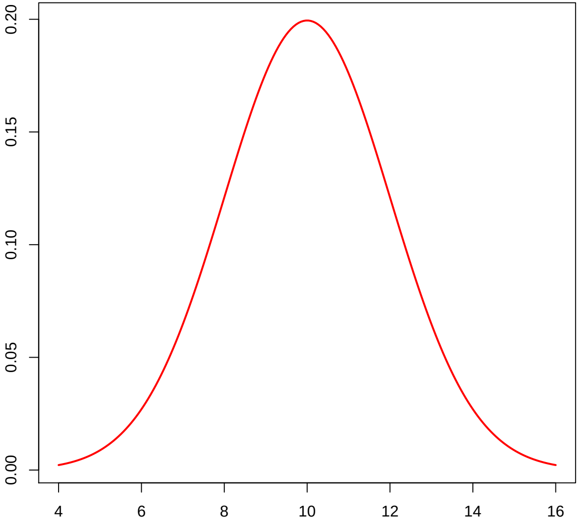
- A general normal distribution has the following probability density function for any given parameters μ and $\sigma \geq 0$:

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}.$$

- The normal distribution is also often denoted as

$$X \sim \mathcal{N}(\mu, \sigma^2).$$

General Normal Random Variable: $\mu = 10$ and $\sigma = 2$



General normal curve: Plot R code

```
> mu<-10
>sigma<-2
>x<-seq(mu-3*sigma,mu+3*sigma,length=200)
>y<-dnorm(x,mean=mu,sd=sigma)
>plot(x,y,type="l",lwd=2,col="red")
```

Properties of Standard Normal Random Variable

- The pdf is symmetric around its mean $x = \mu$, which is at the same time the mode, the median of the distribution.
- It is unimodal.
- It has inflection points at $\mu \pm \sigma$.
- X has mean μ and variance σ; namely

$$\begin{aligned} \mathbb{E}[Z] &= \mu \\ \mathbb{V}[Z] &= \sigma \end{aligned}$$

- **The 68-95-99.7 (empirical) rule, or the 3-sigma rule**: About 68% of values drawn from a normal distribution are within one standard deviation away from the mean; about 95% of the values lie within two standard deviations; and about 99.7% are within three standard deviations.

Layout

1. Continuous Random Variable
   - Probability Density Function (pdf)
   - Probability of any set of real numbers
2. Normal Random Variable
   - Standard Normal Random Variable
   - General Normal Random Variable
3. Relationship between $Z \sim N(0,1)$ and $X \sim N(\mu, \sigma^2)$
4. Calculations with Standard Normal Random Variable via the Normal Table
   - Given z-value, calculate probability
   - Given probability, calculate z-value
5. Calculations with General Normal Random Variable via the Normal Table
   - Given x-value, calculate probability
   - Given probability, calculate x-value

Relationship between $Z \sim N(0,1)$ and $X \sim N(\mu, \sigma^2)$

- Given $X \sim N(\mu, \sigma^2)$, then

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

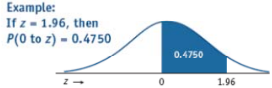
- Given $Z \sim N(0,1)$, then

$$X = \mu + \sigma Z \sim N(\mu, \sigma^2)$$

Layout

The normal table

Areas under the Normal Curve

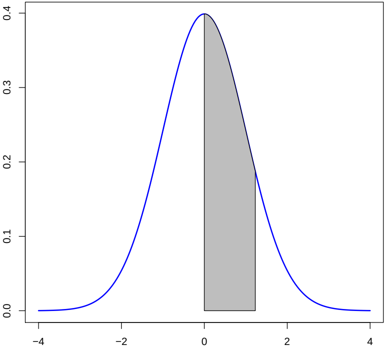


z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.0000	0.0040	0.0080	0.0120	0.0160	0.0199	0.0239	0.0279	0.0319	0.0359
0.1	0.0398	0.0438	0.0478	0.0517	0.0557	0.0596	0.0636	0.0675	0.0714	0.0753
0.2	0.0793	0.0832	0.0871	0.0910	0.0948	0.0987	0.1026	0.1064	0.1103	0.1141
0.3	0.1179	0.1217	0.1255	0.1293	0.1331	0.1368	0.1406	0.1443	0.1480	0.1517
0.4	0.1554	0.1591	0.1628	0.1664	0.1700	0.1736	0.1772	0.1808	0.1844	0.1879
0.5	0.1915	0.1950	0.1985	0.2019	0.2054	0.2088	0.2123	0.2157	0.2190	0.2224
0.6	0.2257	0.2291	0.2324	0.2357	0.2389	0.2422	0.2454	0.2486	0.2517	0.2549
0.7	0.2580	0.2611	0.2642	0.2673	0.2704	0.2734	0.2764	0.2794	0.2823	0.2852
0.8	0.2881	0.2910	0.2939	0.2967	0.2995	0.3023	0.3051	0.3078	0.3106	0.3133
0.9	0.3159	0.3186	0.3212	0.3238	0.3264	0.3289	0.3315	0.3340	0.3365	0.3389
1.0	0.3413	0.3438	0.3461	0.3485	0.3508	0.3531	0.3554	0.3577	0.3599	0.3621
1.1	0.3643	0.3665	0.3686	0.3708	0.3729	0.3749	0.3770	0.3790	0.3810	0.3830
1.2	0.3849	0.3869	0.3888	0.3907	0.3925	0.3944	0.3962	0.3980	0.3997	0.4015
1.3	0.4032	0.4049	0.4066	0.4082	0.4099	0.4115	0.4131	0.4147	0.4162	0.4177
1.4	0.4192	0.4207	0.4222	0.4236	0.4251	0.4265	0.4279	0.4292	0.4306	0.4319
1.5	0.4332	0.4345	0.4357	0.4370	0.4382	0.4394	0.4406	0.4418	0.4429	0.4441
1.6	0.4452	0.4463	0.4474	0.4484	0.4495	0.4505	0.4515	0.4525	0.4535	0.4545
1.7	0.4554	0.4564	0.4573	0.4582	0.4591	0.4599	0.4608	0.4616	0.4625	0.4633
1.8	0.4641	0.4649	0.4656	0.4664	0.4671	0.4678	0.4686	0.4693	0.4699	0.4706
1.9	0.4713	0.4719	0.4726	0.4732	0.4738	0.4744	0.4750	0.4756	0.4761	0.4767
2.0	0.4772	0.4778	0.4783	0.4788	0.4793	0.4798	0.4803	0.4808	0.4812	0.4817
2.1	0.4821	0.4826	0.4830	0.4834	0.4838	0.4842	0.4846	0.4850	0.4854	0.4857
2.2	0.4861	0.4864	0.4868	0.4871	0.4875	0.4878	0.4881	0.4884	0.4887	0.4890
2.3	0.4893	0.4896	0.4898	0.4901	0.4904	0.4906	0.4909	0.4911	0.4913	0.4916
2.4	0.4918	0.4920	0.4922	0.4925	0.4927	0.4929	0.4931	0.4932	0.4934	0.4936
2.5	0.4938	0.4940	0.4941	0.4943	0.4945	0.4946	0.4948	0.4949	0.4951	0.4952
2.6	0.4953	0.4955	0.4956	0.4957	0.4959	0.4960	0.4961	0.4962	0.4963	0.4964
2.7	0.4965	0.4966	0.4967	0.4968	0.4969	0.4970	0.4971	0.4972	0.4973	0.4974
2.8	0.4974	0.4975	0.4976	0.4977	0.4977	0.4978	0.4979	0.4979	0.4980	0.4981

Given z-value, calculate probability

- **Example**: Calculate the area between 0 and 1.23.
- **Solution**: The area is equal to the probability between 0 and 1.23 under the standard normal curve. So from the table

$$P(0 \leq Z \leq 1.23) = 0.3907.$$



R code

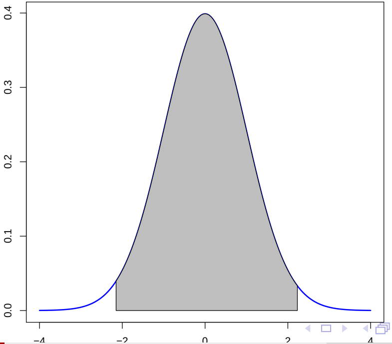
- ```
  > mu<-1
  >sigma<-0
  > pnorm(1.23, mean=mu, sd=sigma)-0.5
  #[1] 0.3906514
  ```
- Or simply run the following code for the standard normal distribution where $\mu = 0$ and $\sigma = 1$

  ```
  > pnorm(1.23)-0.5
  #[1] 0.3906514
  ```
  ```

Given z-value, calculate probability

- **Example**: Calculate the area between -2.15 and 2.23.
- **Solution**: The area is equal to the probability between -2.15 and 1.23 under the standard normal curve. So from the table

$$\begin{aligned} P(-2.15 \le Z \le 2.23) &= P(-2.15 \le Z \le 0) + P(0 \le Z \le 2.23) \\ &= P(0 \le Z \le 2.15) + P(0 \le Z \le 2.23) \\ &= 0.4842 + 0.4871 = 0.9713 \end{aligned}$$



R code

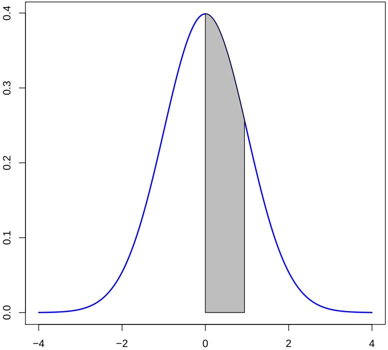
```
> z1<--2.15
> z2<-2.23
> mu<-0
> sigma<-1
> pnorm(z2, mean=mu, sd=sigma)-pnorm(z1, mean=mu, sd=sigma)
[1] 0.9713487
```

Given probability, calculate z-value

- **Example**: Given the area between 0 and z is 0.3264, find z.
- **Solution**: We want to find z such that

$$P(0 \leq Z \leq z) = 0.3264$$

- From the table, we find $z = 0.94$.



R code

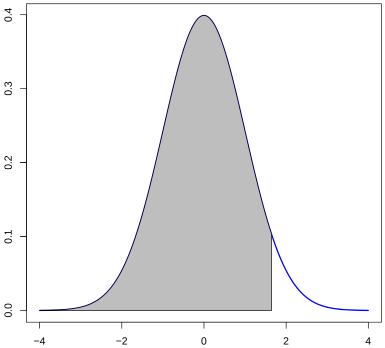
```
>p<-0.3264
>qnorm(p+0.5)
[1] 0.9400342
```

Given probability, calculate z-value

- **Example**: Given the area between less than z is 0.95, find z.
- **Solution**: We want to find z such that

$$P(Z \leq z) = 0.95 \Leftrightarrow P(0 \leq Z \leq z) = 0.95 - 0.5 = 0.45$$

- From the table, we find $z = 1.65$.



R code

```
>p<-0.95
>qnorm(p)
[1] 1.644854
```

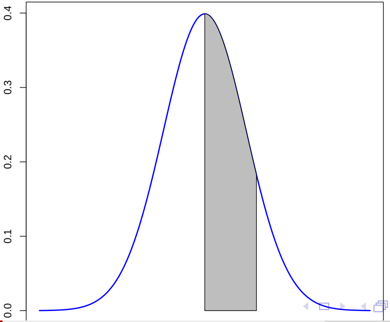
Layout

1. Continuous Random Variable
   - Probability Density Function (pdf)
   - Probability of any set of real numbers
2. Normal Random Variable
   - Standard Normal Random Variable
   - General Normal Random Variable
3. Relationship between $Z \sim N(0,1)$ and $X \sim N(\mu, \sigma^2)$
4. Calculations with Standard Normal Random Variable via the Normal Table
   - Given z-value, calculate probability
   - Given probability, calculate z-value
5. **Calculations with General Normal Random Variable via the Normal Table**
   - Given x-value, calculate probability
   - Given probability, calculate x-value

Given x-value, calculate probability

- **Example**: Given a normal random variable $X \sim N(50, 8^2)$, calculate the area between 50 and 60.
- **Solution**: The area is equal to the probability between 50 and 60 under the normal curve.

$$
\begin{aligned}
P(50 \le X \le 60) &= P\left(\frac{50-50}{8} \le \frac{X-\mu}{\sigma} \le \frac{60-50}{8}\right) \\
&= P(0 \le Z \le 1.25) = 0.394
\end{aligned}
$$



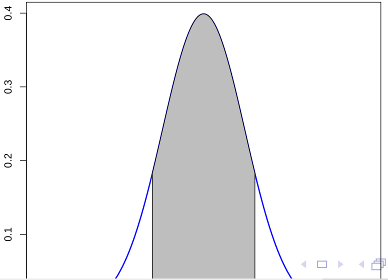
R code

```
> pnorm(1.25)-0.5
[1] 0.3943502
```

Given x-value, calculate probability

- **Example**: Given a normal random variable $X \sim N(50, 8^2)$, calculate the area between 40 and 60.
- **Solution**: The area is equal to the probability between 40 and 60 under the normal curve.

$$\begin{aligned} P(40 \le X \le 60) &= P\left(\frac{40-50}{8} \le \frac{X-\mu}{\sigma} \le \frac{60-50}{8}\right) \\ &= P(-1.25 \le Z \le 1.25) = 2P(0 \le Z \le 1.25) \\ &= 2(0.394) = 0.688 \end{aligned}$$



R code

```
> pnorm(1.25)-pnorm(-1.25)
[1] 0.7887005
```

Given probability, calculate x-value

- **Example**: Given a normal random variable $X \sim N(50, 8^2)$, and the area below x is 0.853, find x?
- **Solution**: We want to find x such that

$$\begin{aligned} P(X \leq x) = 0.853 \quad &\Leftrightarrow \quad P\left(\frac{X-\mu}{\sigma} \leq \frac{x-\mu}{\sigma}\right) = 0.853 \\ &\Leftrightarrow \quad P(Z \leq z) = 0.853 \\ &\Leftrightarrow \quad P(0 \leq Z \leq z) = 0.853 - 0.5 = 0.353, \end{aligned}$$

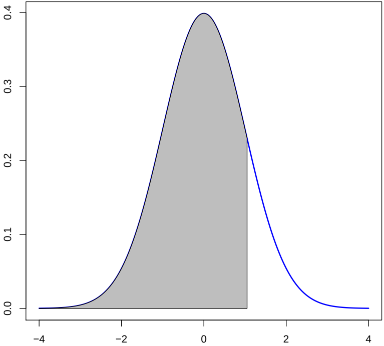
where

$$z := \frac{x-\mu}{\sigma}$$

- From the table, we find $z = 1.05$, implying that

$$x = \mu + z\sigma = 50 + 1.05(8) = 58.4$$

Plot



R code

```
>p<-0.853
>qnorm(p)
[1] 1.049387
```

Practical example

- **Example**: Professor X has determined that the scores in his statistics course are approximately normally distributed with a mean of 72 and a standard deviation of 5. He announces to the class that the top 15 percent of the scores will earn an A.
- **Problem**: What is the lowest score a student can earn and still receive an A?

Practical example

- **Solution**: Let X be the students' scores. Then $X \sim N(72, 5^2)$. Let x be the score that separates an A from the rest. Then

$$\begin{aligned} P(X \geq x) = 0.15 \quad &\Leftrightarrow \quad P\left(\frac{X-\mu}{\sigma} \geq \frac{x-\mu}{\sigma}\right) = 0.15 \\ &\Leftrightarrow \quad P(Z \geq z) = 0.15 \\ &\Leftrightarrow \quad P(0 \leq Z \leq z) = 0.5 - 0.15 = 0.35, \end{aligned}$$

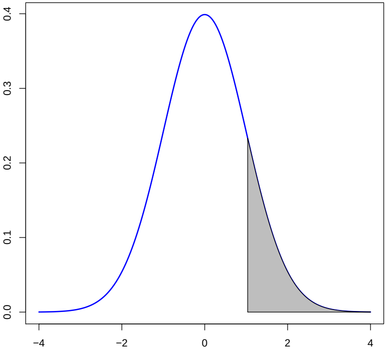
where

$$z := \frac{x-\mu}{\sigma}$$

- From the table, we find $z = 1.04$, implying that

$$x = \mu + z\sigma = 72 + 1.04(5) = 77.2$$

Plot



R code

```
>p<-0.85
>qnorm(p)
[1] 1.036433
```

Practical example

- **Example**: A manufacturer of aircraft is likely to be very concerned about the ability of potential users to use the product. If a lot of pilots cannot reach the rudder pedals or the navigation systems, then there is trouble. Suppose a manufacturer knows that the lengths of pilot's legs are normally distributed with mean 76 and standard deviation of 5 cm.
- **Problem**: If the manufacturer wants to design a cockpit such that precisely 90% of pilots can reach the rudder pedals with their feet while seated, what is the desired distance between seat and pedals?

Practical example

- **Solution**: Let X be the the lengths of pilot's legs. Then $X \sim N(76, 5^2)$. Let x be the desired distance. Then

$$\begin{aligned} P(X \geq x) = 0.90 \quad &\Leftrightarrow \quad P\left(\frac{X-\mu}{\sigma} \geq \frac{x-\mu}{\sigma}\right) = 0.90 \\ &\Leftrightarrow \quad P(Z \geq z) = 0.90 \\ &\Leftrightarrow \quad P(z \leq Z \leq 0) = 0.9 - 0.5 = 0.4 \\ &\Leftrightarrow \quad P(0 \leq Z \leq -z) = 0.4 \end{aligned}$$

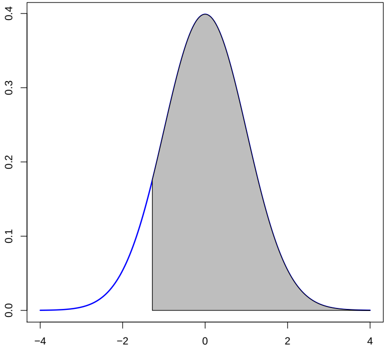
  where

$$z := \frac{x-\mu}{\sigma}$$

- From the table, we find $z = -1.28$, implying that

$$x = \mu + z\sigma = 76 - 1.28(5) = 69.6$$

Plot



R code

```
>p<-0.10
>qnorm(p)
[1] -1.281552
```

Practical example

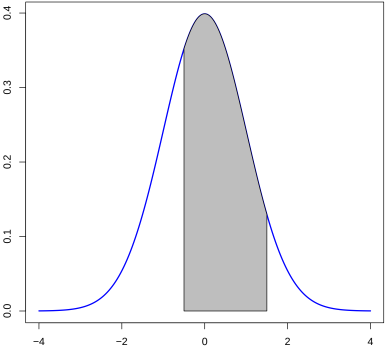
- **Example**: Suppose that a manufacturer of aircraft engines knows their lifetimes to be a normally distributed random variable with a mean of 2,000 hours and a standard deviation of 100 hours
- **Problem**: What is the probability that a randomly chosen engine has a lifetime between 1,950 and 2,150 hours?

Practical example

- **Solution**: Let X be the lifetimes of their aircraft engines. Then $X \sim N(2000, 100^2)$. Then

$$\begin{aligned} & P(1950 \le X \le 2150) \\ = \; & P\left(\frac{1950-2000}{100} \le \frac{X-\mu}{\sigma} \le \frac{2150-2000}{100}\right) \\ = \; & P(-0.5 \le Z \le 1.5) = P(0 \le Z \le 0.5) + P(0 \le Z \le 1.5) \\ = \; & 0.1915 + 0.4332 = 0.6247. \end{aligned}$$

Plot



R code

```
> z1<--0.5
> z2<-1.5
> pnorm(z2)-pnorm(z1)
[1] 0.6246553
```

The Probability of a market Crash

- **Example**: Suppose that the annualized S&P 500 index returns, $\mu \approx 12\%$ and $\sigma \approx 15\%$.
- **Problem**: A negative surprise: on October 19, 1987, the S&P 500 index dropped more than 23% on one day. What is the probability for such a event?

Solution

- **Solution**: Let r denote the daily return, then r is normally distributed with
  - mean
  $$0.12/252 \approx 0.00048,$$
  - and standard deviation
  $$0.15/\sqrt{252} = 0.0094.$$

  Namely $r \sim N(0.00048, 0.0094^2)$. Then

$$\begin{aligned} & P(r \leq -0.23) \\ = \;& P\left(\frac{r-\mu}{\sigma} \leq \frac{-0.23 - 0.00048}{0.0094}\right) \\ = \;& P(Z \leq -24) \approx 10^{-127}. \end{aligned}$$

The empirical rule

- We now derive the empirical rule (Back in Lecture 4) from the Normal table: assume $X \sim N(\mu, \sigma^2)$, then

$$\mathbb{P}\left(\mu - k\sigma \leq X \leq \mu - k\sigma\right) = \mathbb{P}\left(-k \leq \frac{X-\mu}{\sigma} \leq k\right) = \mathbb{P}(-k \leq Z \leq k)$$

- For $k = 1, 2, 3$, we obtain 0.68, 0.95, and 99.7 from the Normal table.