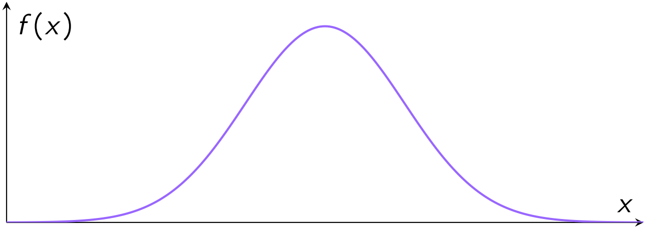
Chapter 5 - Part 2

The Normal Distribution

The Normal Distribution

- The normal distribution is the most important one.
- Also known as: Gaussian distribution or bell-shaped distribution or bell curve.



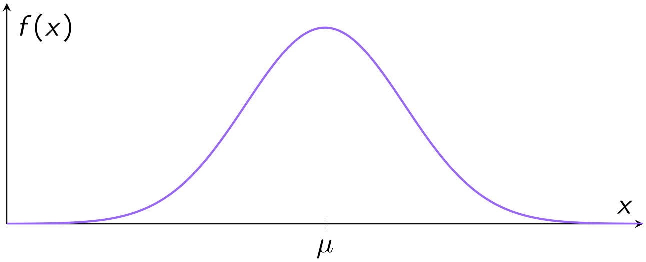
- It plays a very important role in the science of statistical inference.
- Many phenomena generate random variables with probability distributions that are very well approximated by a normal distribution.

The Normal Distribution

A continuous random variable X is said to have a normal distribution if it has a PDF of the form

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

for some numbers μ and $\sigma > 0$.



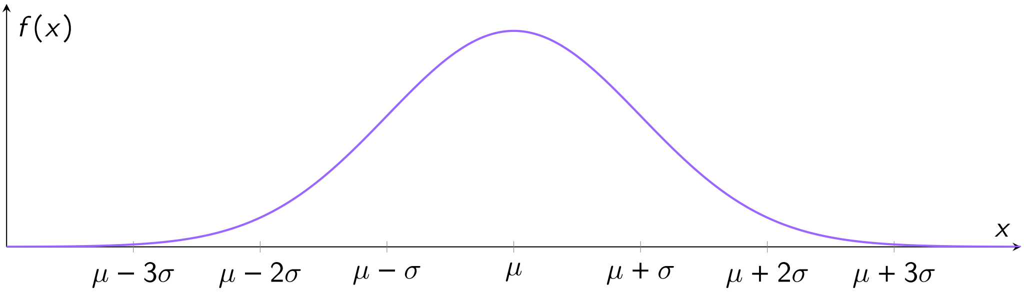
In that case, we denote

$$X \sim N(\mu, \sigma).$$

Expected Value and Variance

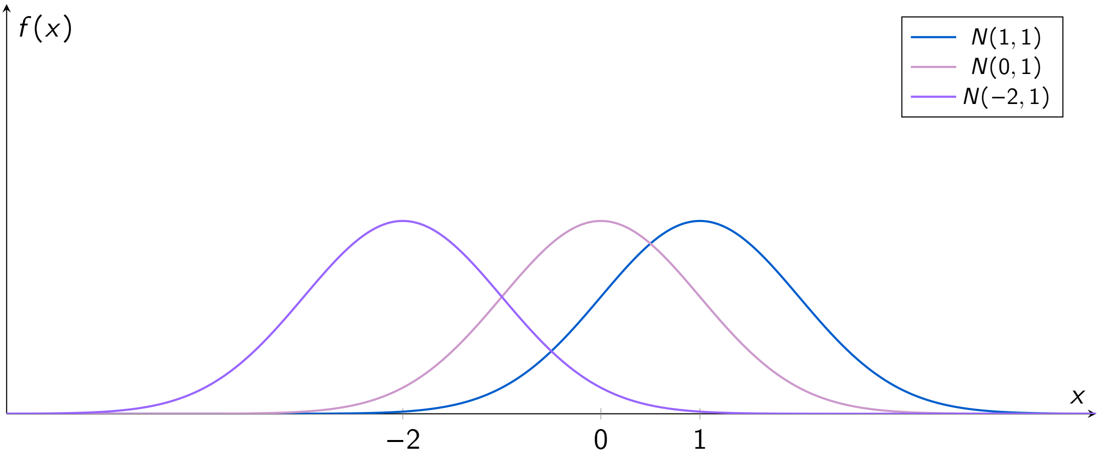
If $X \sim N(\mu, \sigma)$, then

$$E(X) = \mu$$
$$\mathrm{StDev}(X) = \sigma$$



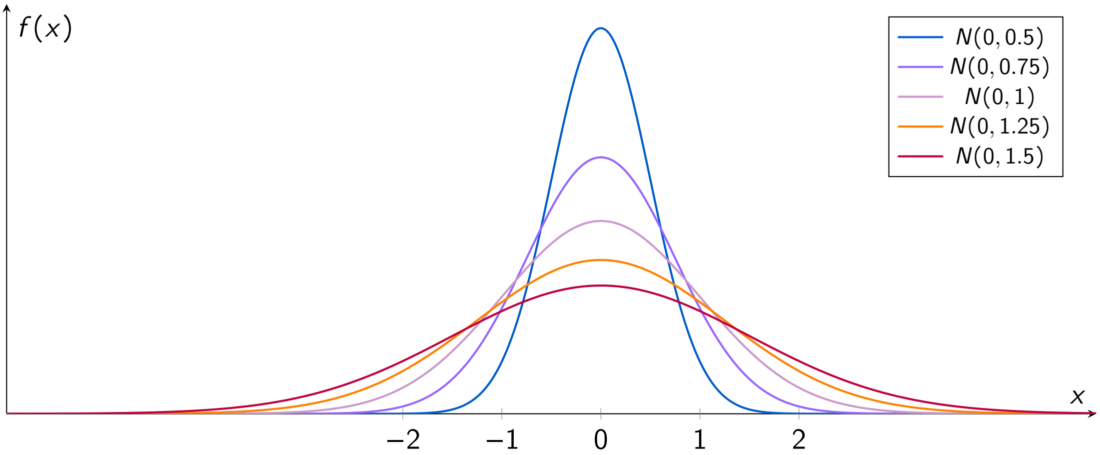
$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

Comparing μ



Change of $N(\mu, \sigma)$ as μ changes

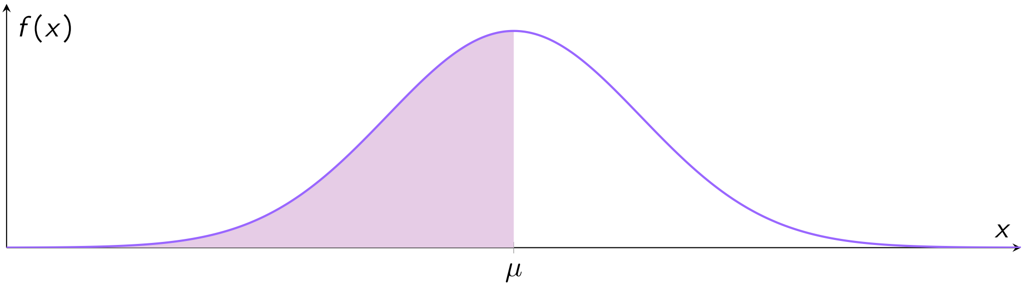
Comparing σ



Change of $N(\mu, \sigma)$ as σ changes

Median

Let $X \sim N(\mu, \sigma)$. Then the median of X is μ, by symmetry.

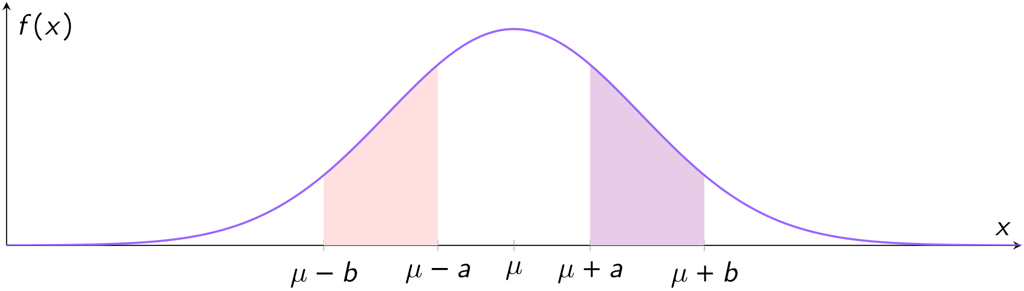


$$P(X < \mu) = P(\mu < X) = 0.5$$

Symmetry

Let $0 < a < b$. Again by symmetry,

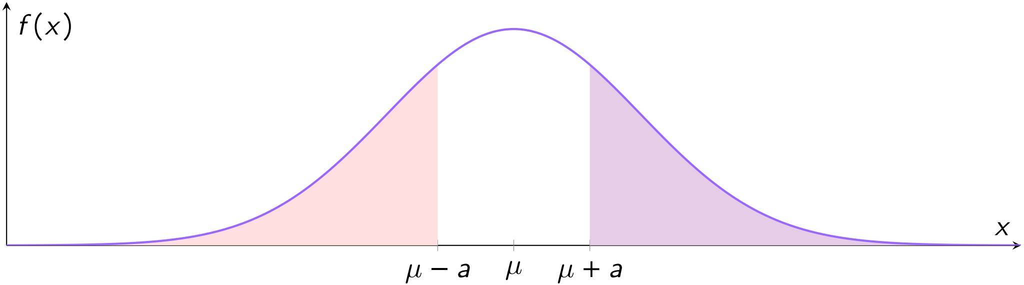
$$P(\mu - b < X < \mu - a) = P(\mu + a < X < \mu + b)$$



Symmetry

Similarly,

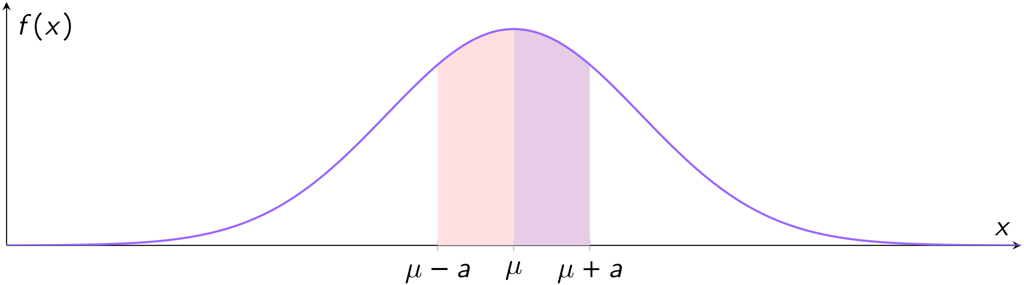
$$P(X < \mu - a) = P(\mu + a < X)$$



Symmetry

and,

$$P(\mu - a < X < \mu) = P(\mu < X < \mu + a)$$

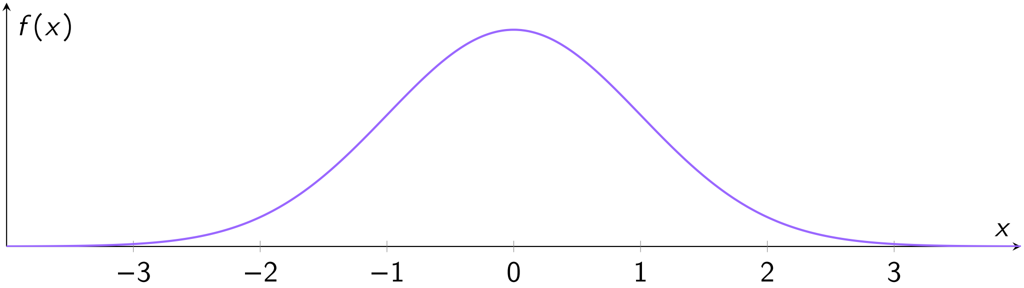


The Standard Normal Distribution

$N(0,1)$ is called the standard normal distribution.

- It is the normal distribution with

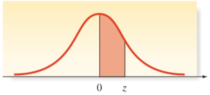
$$\mu = 0, \qquad \sigma = 1$$



$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

For the rest of this course, we will denote $Z \sim N(0,1)$.

The Standard Normal Table



z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
.0	.0000	.0040	.0080	.0120	.0160	.0199	.0239	.0279	.0319	.0359
.1	.0398	.0438	.0478	.0517	.0557	.0596	.0636	.0675	.0714	.0753
.2	.0793	.0832	.0871	.0910	.0948	.0987	.1026	.1064	.1103	.1141
.3	.1179	.1217	.1255	.1293	.1331	.1368	.1406	.1443	.1480	.1517
.4	.1554	.1591	.1628	.1664	.1700	.1736	.1772	.1808	.1844	.1879
.5	.1915	.1950	.1985	.2019	.2054	.2088	.2123	.2157	.2190	.2224
.6	.2257	.2291	.2324	.2357	.2389	.2422	.2454	.2486	.2517	.2549
.7	.2580	.2611	.2642	.2673	.2704	.2734	.2764	.2794	.2823	.2852
.8	.2881	.2910	.2939	.2967	.2995	.3023	.3051	.3078	.3106	.3133
.9	.3159	.3186	.3212	.3238	.3264	.3289	.3315	.3340	.3365	.3389
1.0	.3413	.3438	.3461	.3485	.3508	.3531	.3554	.3577	.3599	.3621
1.1	.3643	.3665	.3686	.3708	.3729	.3749	.3770	.3790	.3810	.3830
1.2	.3849	.3869	.3888	.3907	.3925	.3944	.3962	.3980	.3997	.4015
1.3	.4032	.4049	.4066	.4082	.4099	.4115	.4131	.4147	.4162	.4177
1.4	.4192	.4207	.4222	.4236	.4251	.4265	.4279	.4292	.4306	.4319
1.5	.4332	.4345	.4357	.4370	.4382	.4394	.4406	.4418	.4429	.4441
1.6	.4452	.4463	.4474	.4484	.4495	.4505	.4515	.4525	.4535	.4545
1.7	.4554	.4564	.4573	.4582	.4591	.4599	.4608	.4616	.4625	.4633
1.8	.4641	.4649	.4656	.4664	.4671	.4678	.4686	.4693	.4699	.4706
1.9	.4713	.4719	.4726	.4732	.4738	.4744	.4750	.4756	.4761	.4767
2.0	.4772	.4778	.4783	.4788	.4793	.4798	.4803	.4808	.4812	.4817
2.1	.4821	.4826	.4830	.4834	.4838	.4842	.4846	.4850	.4854	.4857
2.2	.4861	.4864	.4868	.4871	.4875	.4878	.4881	.4884	.4887	.4890
2.3	.4893	.4896	.4898	.4901	.4904	.4906	.4909	.4911	.4913	.4916
2.4	.4918	.4920	.4922	.4925	.4927	.4929	.4931	.4932	.4934	.4936
2.5	.4938	.4940	.4941	.4943	.4945	.4946	.4948	.4949	.4951	.4952
2.6	.4953	.4955	.4956	.4957	.4959	.4960	.4961	.4962	.4963	.4964
2.7	.4965	.4966	.4967	.4968	.4969	.4970	.4971	.4972	.4973	.4974
2.8	.4974	.4975	.4976	.4977	.4977	.4978	.4979	.4979	.4980	.4981
2.9	.4981	.4982	.4982	.4983	.4984	.4984	.4985	.4985	.4986	.4986
3.0	.4987	.4987	.4987	.4988	.4988	.4989	.4989	.4989	.4990	.4990

The Standard Normal Table

We will read the probability values $P(0 < Z < z)$ from the standard normal table.

Example: For $P(0 < Z < 1.19)$ we will look at $z = 1.19$ from the table.

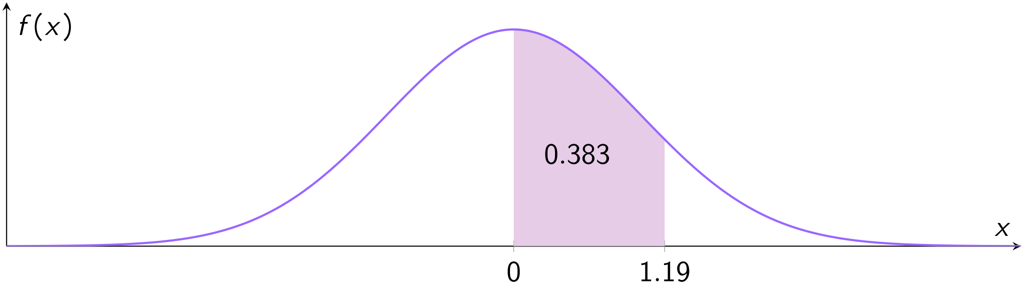
z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
.0	.0000	.0040	.0080	.0120	.0160	.0199	.0239	.0279	.0319	.0359
.1	.0398	.0438	.0478	.0517	.0557	.0596	.0636	.0675	.0714	.0753
.2	.0793	.0832	.0871	.0910	.0948	.0987	.1026	.1064	.1103	.1141
.3	.1179	.1217	.1255	.1293	.1331	.1368	.1406	.1443	.1480	.1517
.4	.1554	.1591	.1628	.1664	.1700	.1736	.1772	.1808	.1844	.1879
.5	.1915	.1950	.1985	.2019	.2054	.2088	.2123	.2157	.2190	.2224
.6	.2257	.2291	.2324	.2357	.2389	.2422	.2454	.2486	.2517	.2549
.7	.2580	.2611	.2642	.2673	.2704	.2734	.2764	.2794	.2823	.2852
.8	.2881	.2910	.2939	.2967	.2995	.3023	.3051	.3078	.3106	.3133
.9	.3159	.3186	.3212	.3238	.3264	.3289	.3315	.3340	.3365	.3389
1.0	.3413	.3438	.3461	.3485	.3508	.3531	.3554	.3577	.3599	.3621
1.1	.3643	.3665	.3686	.3708	.3729	.3749	.3770	.3790	.3810	.3830
1.2	.3849	.3869	.3888	.3907	.3925	.3944	.3962	.3980	.3997	.4015
1.3	.4032	.4049	.4066	.4082	.4099	.4115	.4131	.4147	.4162	.4177
1.4	.4192	.4207	.4222	.4236	.4251	.4265	.4279	.4292	.4306	.4319
1.5	.4332	.4345	.4357	.4370	.4382	.4394	.4406	.4418	.4429	.4441

$$z = 1.19 = 1.1 + 0.09$$

This means that

$$P(0 < Z < 1.19) = 0.383$$

The Standard Normal Table



$$P(0 < Z < 1.19) = 0.383$$

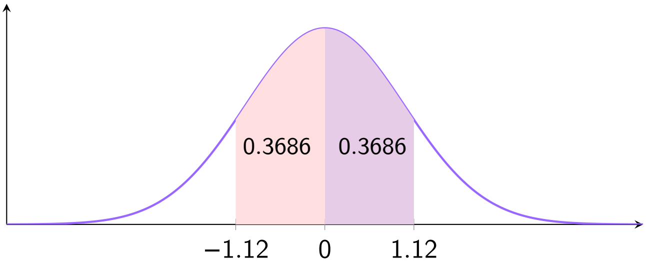
Note: The numbers in the table are integral values. For example:

$$\int_0^{1.19} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}\, dx \approx 0.383$$

Examples

$Z \sim N(0,1)$

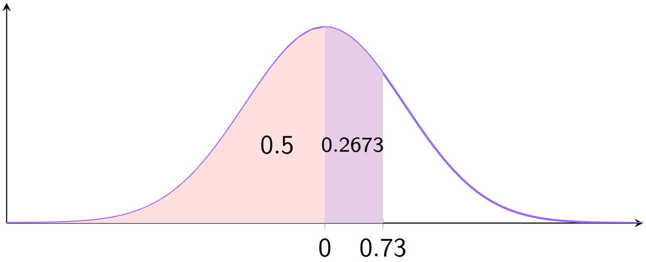
- $P(0 < Z) = 0.5$
- $P(Z < 0) = 0.5$
- $P(0 < Z < 0.62) = 0.2324$
- $P(-1.12 < Z < 0) = P(0 < Z < 1.12) = 0.3686$



Examples

$Z \sim N(0, 1)$

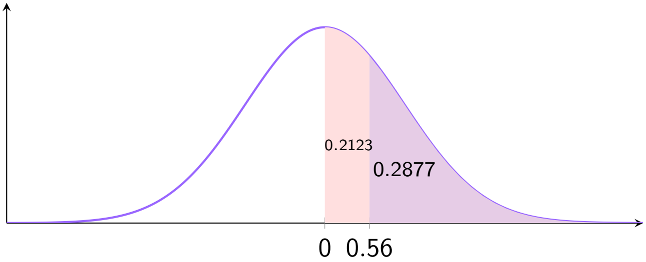
- $P(Z < 0.73) = \underbrace{P(Z < 0)}_{0.5} + \underbrace{P(0 < Z < 0.73)}_{0.2673} = 0.7673$



Examples

$Z \sim N(0,1)$

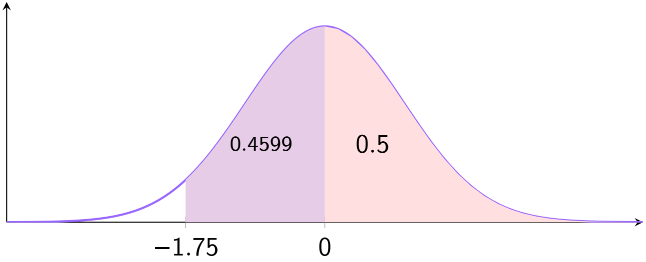
- $P(0.56 < Z) = \underbrace{P(0 < Z)}_{0.5} - \underbrace{P(0 < Z < 0.56)}_{0.2123} = 0.2877$



Examples

$Z \sim N(0,1)$

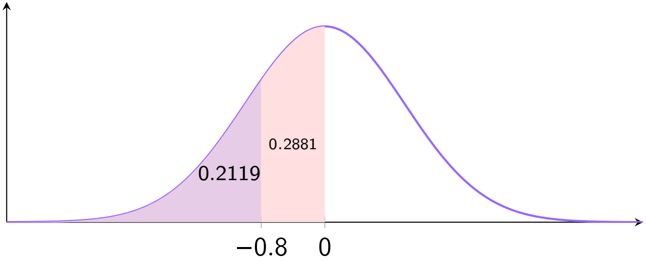
- $P(-1.75 < Z) = \underbrace{P(-1.75 < Z < 0)}_{P(0<Z<1.75)=0.4599} + \underbrace{P(0 < Z)}_{0.5} = 0.9599$



Examples

$Z \sim N(0, 1)$

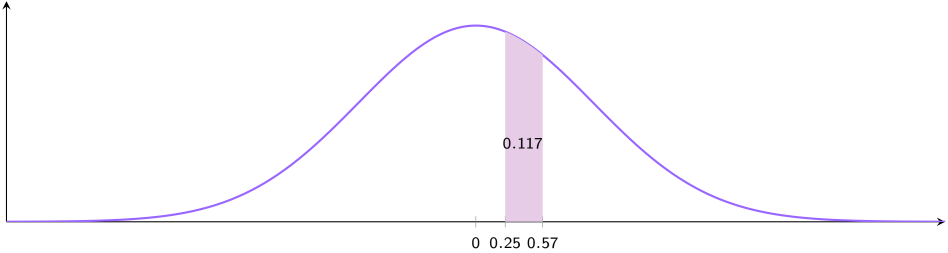
- $P(Z < -0.8) = \underbrace{P(Z < 0)}_{0.5} - \underbrace{P(-0.8 < Z < 0)}_{P(0<Z<0.8)=0.2881} = 0.2119$



Examples

$Z \sim N(0,1)$

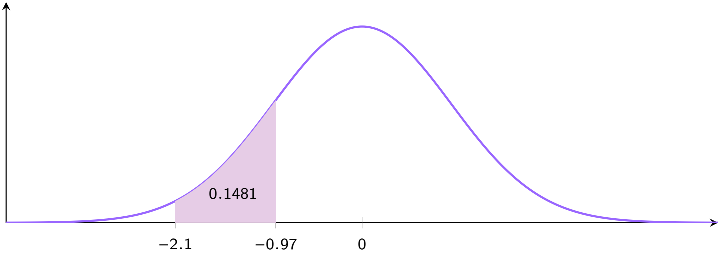
- $P(0.25 < Z < 0.57) = \underbrace{P(0 < Z < 0.57)}_{0.2157} - \underbrace{P(0 < Z < 0.25)}_{0.0987} = 0.117$



Examples

$Z \sim N(0,1)$

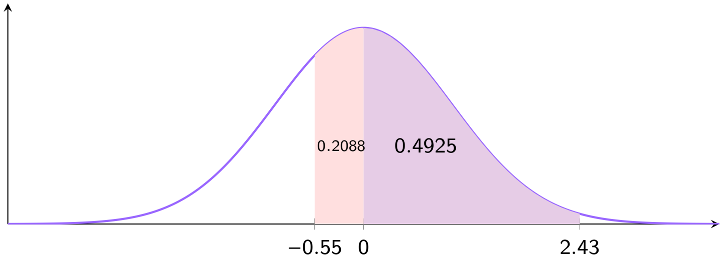
- $P(-2.1 < Z < -0.97) = P(0.97 < Z < 2.1) = \underbrace{P(0 < Z < 2.1)}_{0.4821} - \underbrace{P(0 < Z < 0.97)}_{0.334} = 0.1481$



Examples

$Z \sim N(0,1)$

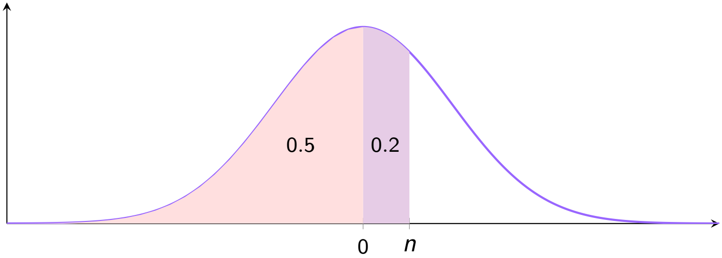
- $P(-0.55 < Z < 2.43) = \underbrace{P(-0.55 < Z < 0)}_{P(0<Z<0.55)=0.2088} + \underbrace{P(0 < Z < 2.43)}_{0.4925} = 0.7013$



Examples

$Z \sim N(0,1)$

- Find the 70$^{\text{th}}$ percentile of Z.
  Solution: We look for n such that $P(Z < n) = 0.7$. Since $0.7 > 0.5$, we know that $n > 0$.



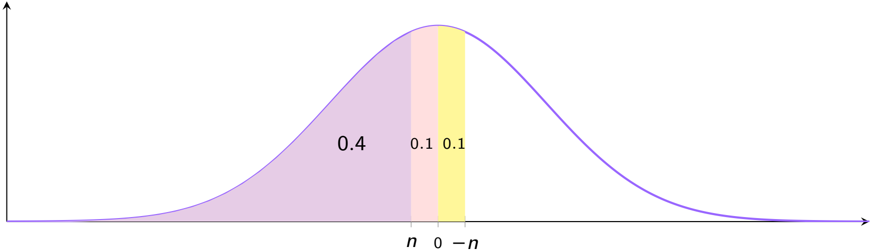
$$0.7 = P(Z < n) = \underbrace{P(Z < 0)}_{0.5} + P(0 < Z < n) \implies P(0 < Z < n) = 0.2 \implies n \approx 0.52$$

Examples

$Z \sim N(0, 1)$

- Find the 40^{th} percentile of Z.

  Solution: We look for n such that $P(Z < n) = 0.4$. Since $0.4 < 0.5$, we know that $n < 0$.



$$0.4 = P(Z < n) = \underbrace{P(Z < 0)}_{0.5} - P(n < Z < 0) \implies P(n < Z < 0) = 0.1$$

$$\implies P(0 < Z < -n) = 0.1$$

$$\implies -n \approx 0.25 \implies n \approx -0.25$$

Converting a Normal Distribution to a Standard Normal Distribution

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

Proof.

If $X \sim N(\mu, \sigma)$ is normal, then $Z = \frac{X-\mu}{\sigma}$ is also normal. Moreover,

$$E(Z) = \frac{E(X) - \mu}{\sigma} = \frac{\mu - \mu}{\sigma} = 0$$

and

$$\mathrm{StDev}(Z) = \frac{\mathrm{StDev}(X)}{\sigma} = \frac{\sigma}{\sigma} = 1.$$

Hence $Z \sim N(0,1)$. $\square$

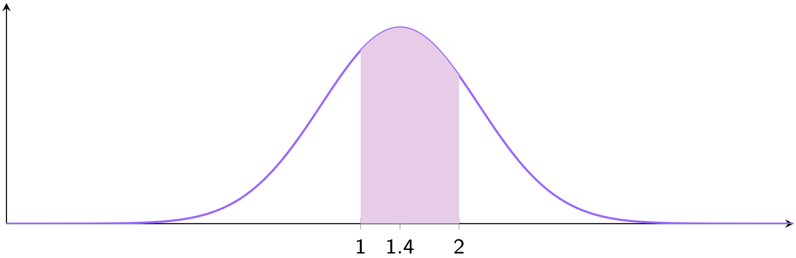
Converting a Normal Distribution to a Standard Normal Distribution

Therefore: Let $Z \sim N(0, 1)$. If $X \sim N(\mu, \sigma)$,

$$P(a < X < b) = P\left(\frac{a - \mu}{\sigma} < Z < \frac{b - \mu}{\sigma}\right).$$

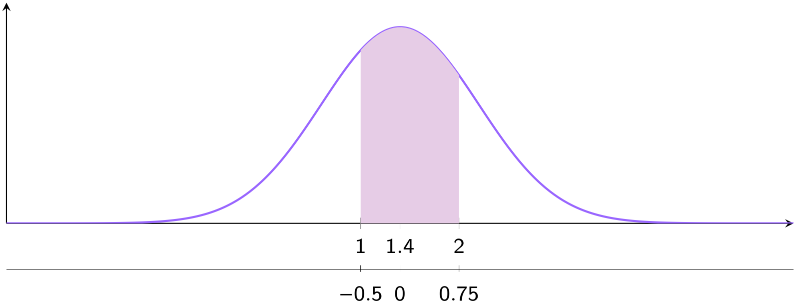
Example

Let $X \sim N(1.4, 0.8)$. Find $P(1 < X < 2)$.



Example - Solution

Let $X \sim N(1.4, 0.8)$. Find $P(1 < X < 2)$.

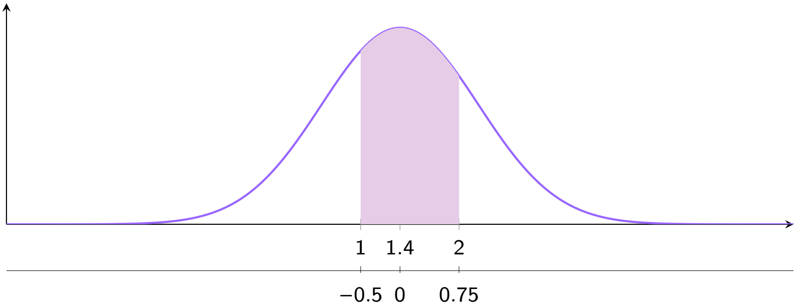


$$\frac{1-\mu}{\sigma} = \frac{1-1.4}{0.8} = -0.5, \qquad \frac{2-\mu}{\sigma} = \frac{2-1.4}{0.8} = 0.75$$

Hence

$$P(1 < X < 2) = P(-0.5 < Z < 0.75)$$

Example - Solution (cont'd)



$$
\begin{aligned}
P(1 < X < 2) &= P(-0.5 < Z < 0.75) \\
&= P(-0.5 < Z < 0) + P(0 < Z < 0.75) \\
&= P(0 < Z < 0.5) + P(0 < Z < 0.75) \\
&= 0.1915 + 0.2734 \\
&= 0.4649
\end{aligned}
$$

Example

Let $X \sim N(-1, 2)$. Find $P(-1.2 < X)$.

Solution: Since

$$\frac{-1.2 - \mu}{\sigma} = \frac{-1.2 - (-1)}{2} = -0.1$$

we have

$$\begin{aligned} P(-1.2 < X) &= P(-0.1 < Z) \\ &= P(-0.1 < Z < 0) + P(0 < Z) \\ &= P(0 < Z < 0.1) + P(0 < Z) \\ &= 0.0398 + 0.5 \\ &= 0.5398 \end{aligned}$$

Example

Professor Snape teaches a course on "Potions" at Hogwarts. The scores of students in the course follow a normal distribution with a mean of 48 and a standard deviation of 16. If the passing grade is set at 40, what percentage of students will pass the course?

Solution: We look for $P(40 < X)$ where $X \sim N(48, 16)$. Since

$$\frac{40 - 48}{16} = 0.5$$

we have

$$\begin{aligned} P(40 < X) &= P(-0.5 < Z) \\ &= P(-0.5 < Z < 0) + P(0 < Z) \\ &= P(0 < Z < 0.5) + P(0 < Z) \\ &= 0.1915 + 0.5 \\ &= 0.6915 = 69.15\% \end{aligned}$$

Example

Let $X \sim N(0.8, 2.4)$. Find the 75^{th} percentile of X.

Solution: First, we find the 75^{th} percentile of $Z \sim N(0,1)$.

$$P(Z < n) = 0.75 \implies n > 0.$$

Therefore

$$0.75 = P(Z < n) = \underbrace{P(Z < 0)}_{0.5} + P(0 < Z < n) \implies P(0 < Z < n) = 0.25 \implies n \approx 0.67$$

We look for m where $P(X < m) = 0.75$.

$$P(X < m) = 0.75 = P(Z < 0.67) \implies \frac{m - \mu}{\sigma} = 0.67 \implies \frac{m - 0.8}{2.4} = 0.67$$
$$\implies m = 0.8 + 2.4 \cdot 0.67 \approx 2.4$$

Example

Let $X \sim N(\mu, \sigma)$. If $P(X < 1) = 0.62$ and $P(X < -1) = 0.34$, find μ and σ.

Solution: Let

$$z_1 = \frac{-1-\mu}{\sigma}, \quad z_2 = \frac{1-\mu}{\sigma}$$

$$P(Z < z_2) = 0.62 \implies P(0 < Z < z_2) = 0.12 \implies z_2 = 0.31$$
$$P(Z < z_1) = 0.34 \implies P(z_1 < Z < 0) = 0.16 \implies z_1 = -0.41$$

Hence

$$\frac{-1-\mu}{\sigma} = -0.41 \implies -1 = \mu - 0.41\sigma$$
$$\frac{1-\mu}{\sigma} = 0.31 \implies 1 = \mu + 0.31\sigma$$

$$\implies 2 = 0.72\sigma \implies \sigma = \frac{2}{0.72} \approx 2.78$$
$$\implies \mu = 1 - 0.31\sigma \approx 0.14$$