Probability and Statistic (for CS Students)

Lecture 14: Normal Distribution

University of Mazandaran

Fall 1405

Review

The probability density function (PDF) of a continuous random variable represents the derivative of probability at a given point.

Units of probability divided by units of X.
Integrate it to get probabilities!

$$P\left(a < X < b\right) = \int_{x=a}^{b} f\left(X = x\right) dx$$

What do you get if you integrate over a probability density function?

What do you get if you integrate over a probability density function?

A Probability

Uniform Distribution

Notation: $X \sim \text{Uni}(\alpha, \beta)$

Description: A continuous random variable that takes on values, with equal likelihood, between α and β

Parameters: $\alpha \in \mathbb{R}$, the minimum value of the variable.
$\beta \in \mathbb{R}$, $\beta > \alpha$, the maximum value of the variable.

Support: $x \in [\alpha, \beta]$

PDF equation:
$$f(x) = \begin{cases} \frac{1}{\beta - \alpha} & \text{for } x \in [\alpha, \beta] \\ 0 & \text{else} \end{cases}$$

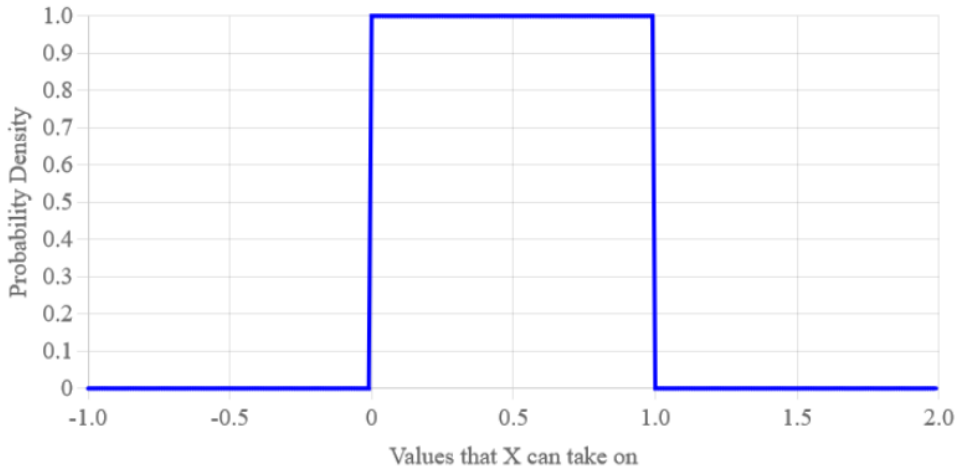
CDF equation:
$$F(x) = \begin{cases} \frac{x - \alpha}{\beta - \alpha} & \text{for } x \in [\alpha, \beta] \\ 0 & \text{for } x < \alpha \\ 1 & \text{for } x > \beta \end{cases}$$

Expectation: $\text{E}[X] = \frac{1}{2}(\alpha + \beta)$

Variance: $\text{Var}(X) = \frac{1}{12}(\beta - \alpha)^2$

PDF graph:

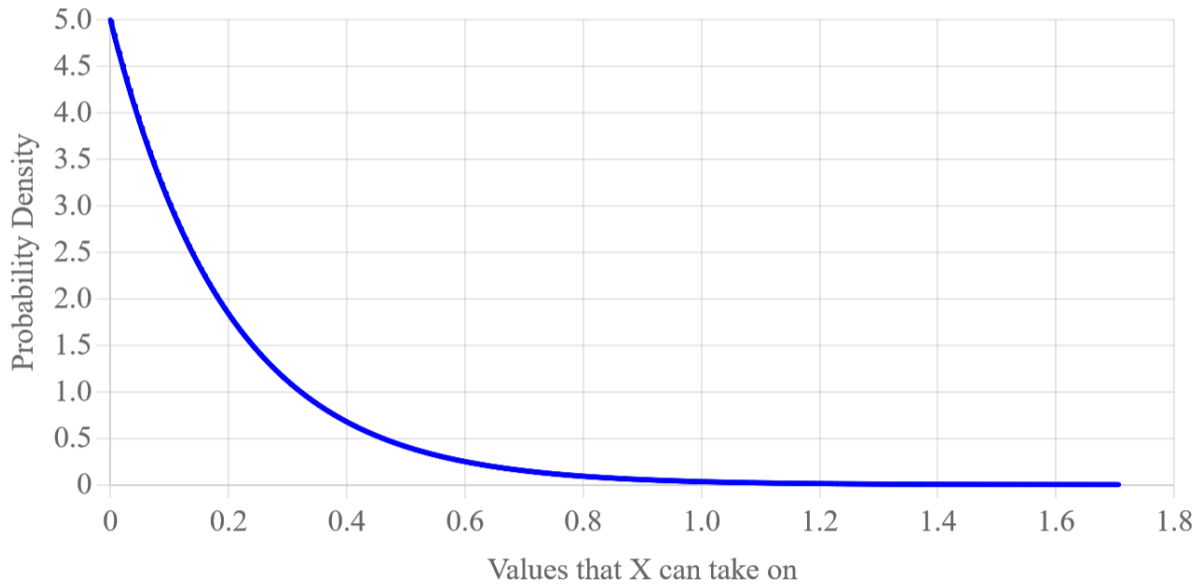
Parameter α: 0 Parameter β: 1



Review

Notation:	$X \sim \text{Exp}(\lambda)$
Description:	Time until next events if (a) the events occur with a constant mean rate and (b) they occur independently of time since last event.
Parameters:	$\lambda \in \mathbb{R}^+$, the constant average rate.
Support:	$x \in \mathbb{R}^+$
PDF equation:	$f(x) = \lambda e^{-\lambda x}$
CDF equation:	$F(x) = 1 - e^{-\lambda x}$
Expectation:	$\text{E}[X] = 1/\lambda$
Variance:	$\text{Var}(X) = 1/\lambda^2$
PDF graph:	

Parameter λ: 5



Review

A cumulative density function (CDF) is a "closed form" equation for the probability that a random variable is less than a given value.

$$F(x) = P(X < x)$$

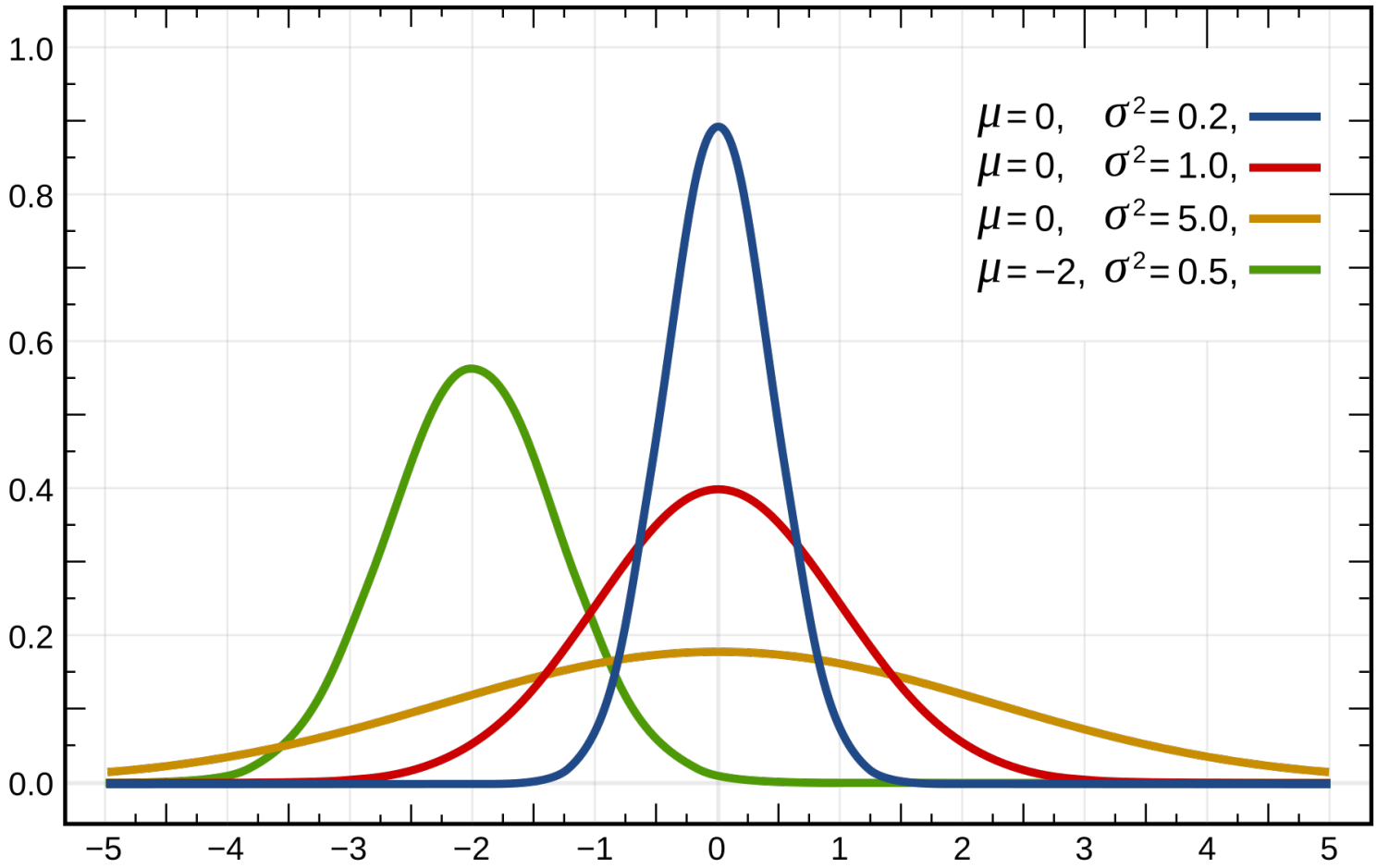
If you learn how to use a cumulative density function, you can avoid integrals!

Let's play a game!

Rules of the game:

- Students must not communicate with one another
- Students must not search for the answer
- Each student submits only one estimate without revealing his/her answer

A Few Normal Examples!



Normal Random Variable

Def A Normal random variable X is defined as follows:

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

Support: $(-\infty, \infty)$

Normal Random Variable

Def A Normal random variable X is defined as follows:

$$X \sim \mathcal{N}\left(\mu, \sigma^2\right)$$
$$\text{Support: } (-\infty, \infty)$$

Other names: **Gaussian** random variable

$$\text{Expectation:} \quad E[X] = \mu$$
$$\text{Variance:} \quad \text{Var}(X) = \sigma^2$$

Normal Random Variable

Def A Normal random variable X is defined as follows:

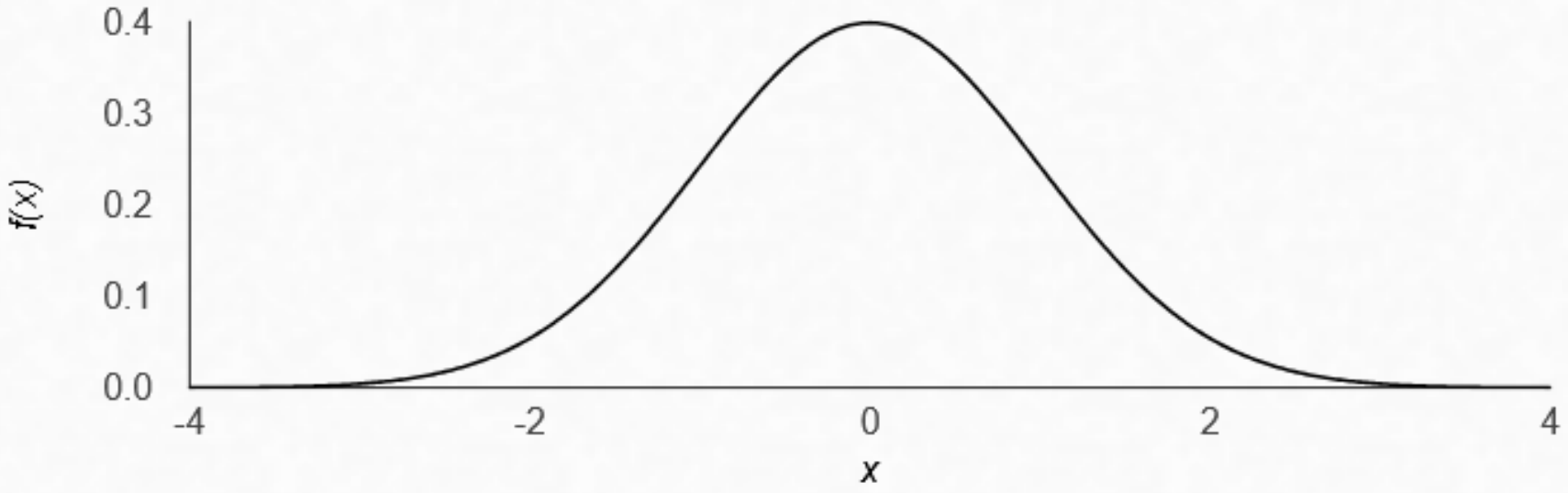
$$X \sim \mathcal{N}\left(\mu, \sigma^2\right)$$
$$\text{Support: } (-\infty, \infty)$$

Other names: **Gaussian** random variable

$$\text{Expectation:} \quad E[X] = \mu$$
$$\text{Variance:} \quad \text{Var}(X) = \sigma^2$$

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

Normal Random Variable



$$\mu = E(X) = 0 \quad \sigma = SD(X) = 1 \quad \sigma^2 = Var(X) = 1$$

Why the Normal?

- Common for natural phenomena: height, weight, etc.
- Most noise in the world in Normal
- Often results from the sum of many random variables
- Sample means are distributed normally

Why the Normal?

- ~~Common for natural phenomena: height, weight, etc.~~

  These are log-normal
- Most noise in the world in Normal
- Often results from the sum of many random variables
- Sample means are distributed normally

Why the Normal?

- ~~Common for natural phenomena: height, weight, etc.~~

  These are log-normal
- ~~Most noise in the world in Normal~~

  Most noise is assumed to be normal!
- Often results from the sum of many random variables
- Sample means are distributed normally

Why the Normal?

- ~~Common for natural phenomena: height, weight, etc.~~

  These are log-normal
- ~~Most noise in the world in Normal~~

  Most noise is assumed to be normal!
- ~~Often results from the sum of many random variables~~

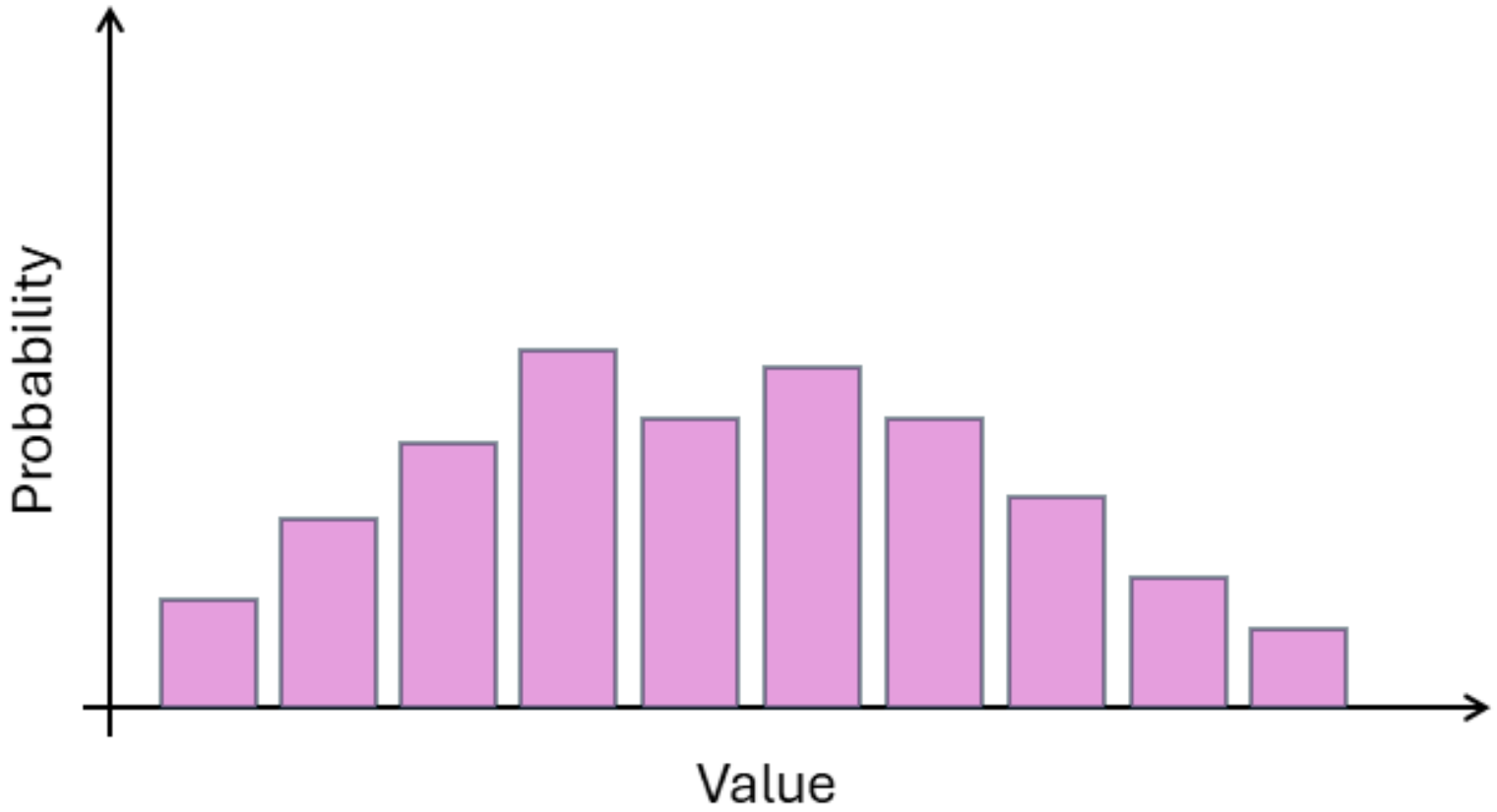
  Only if they are equally weighted and independent
- Sample means are distributed normally

Why the Normal?

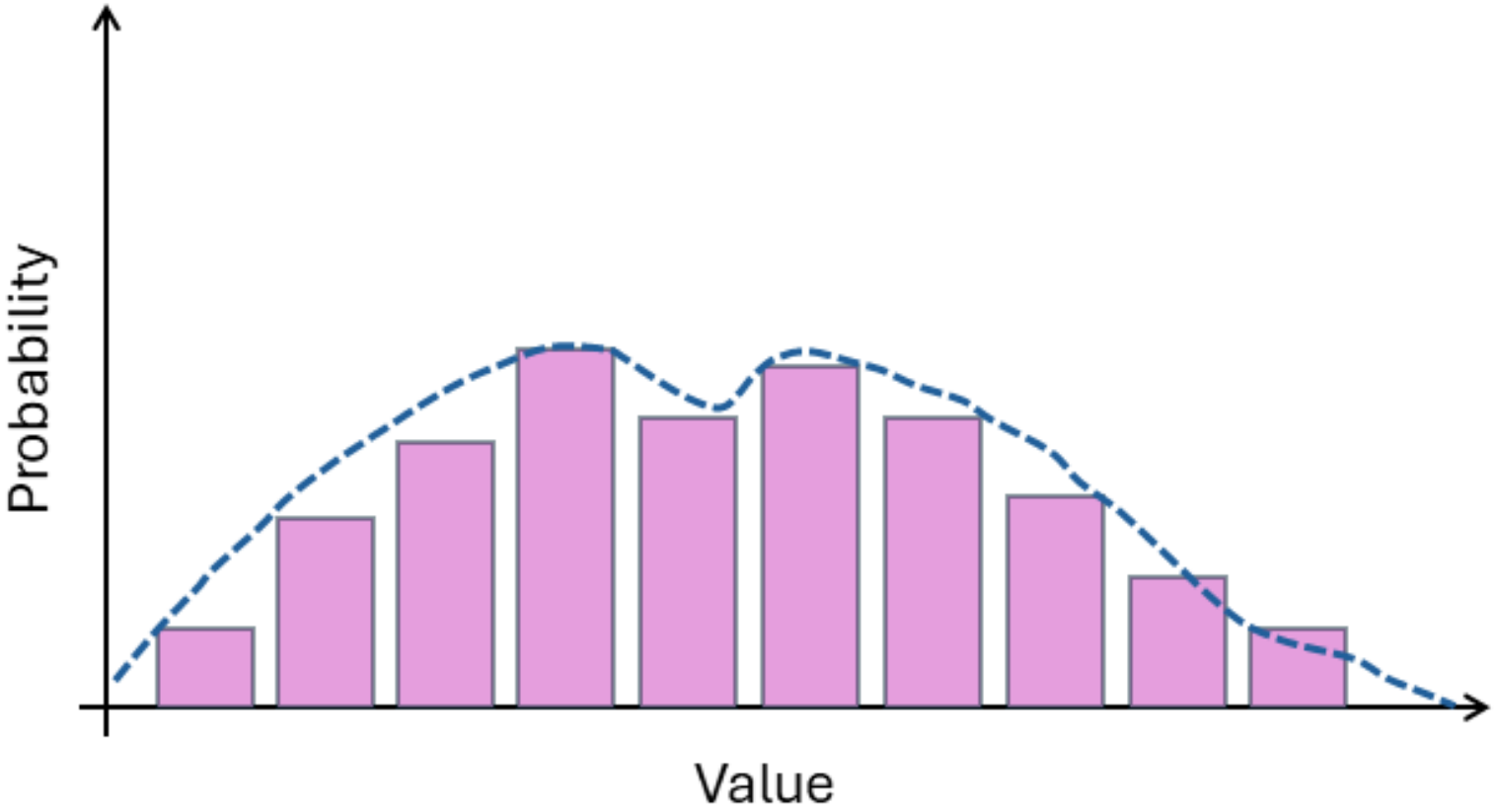
"The simplest explanation is usually the best one!"



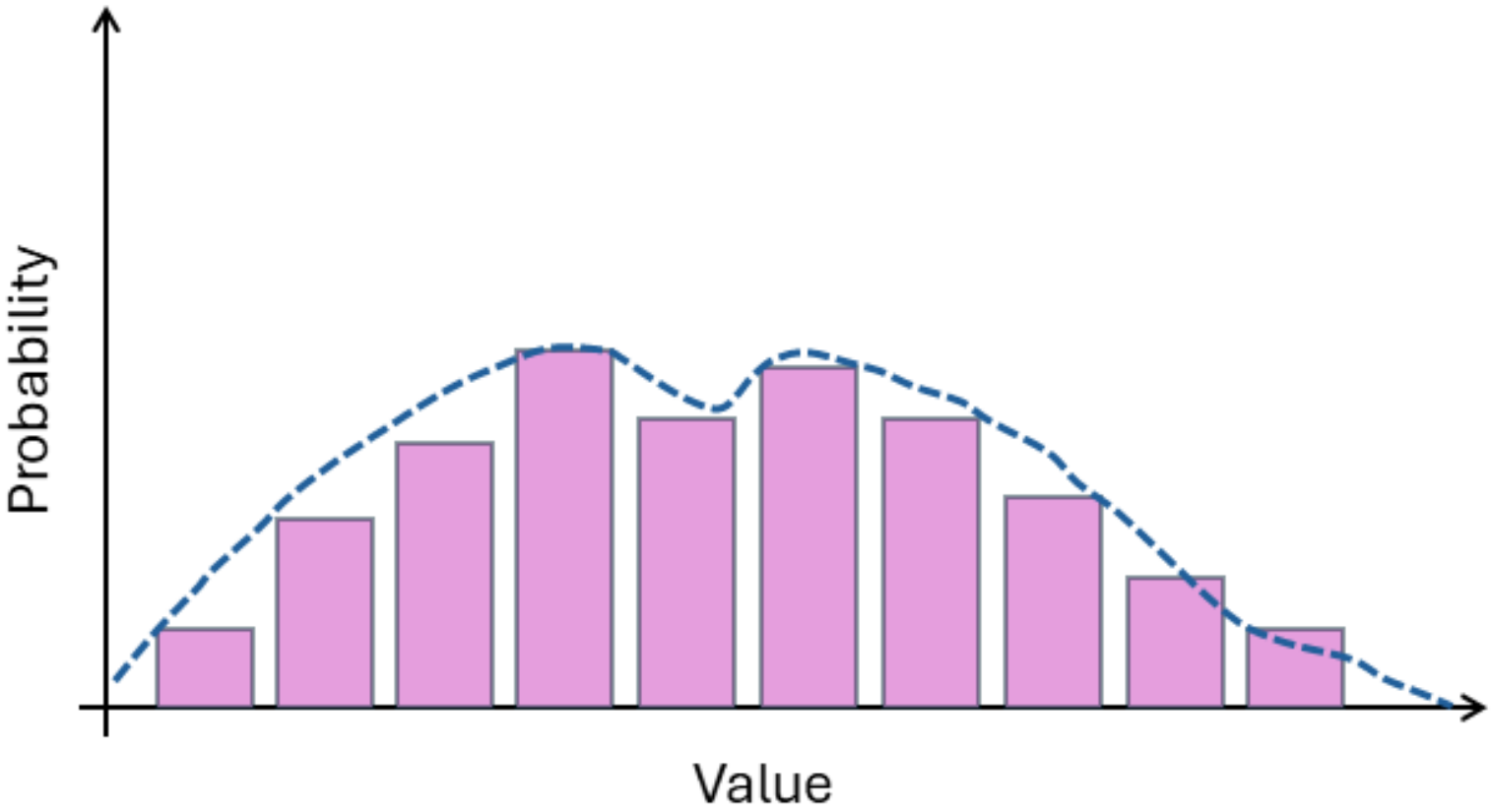
Why the Normal?



Why the Normal?

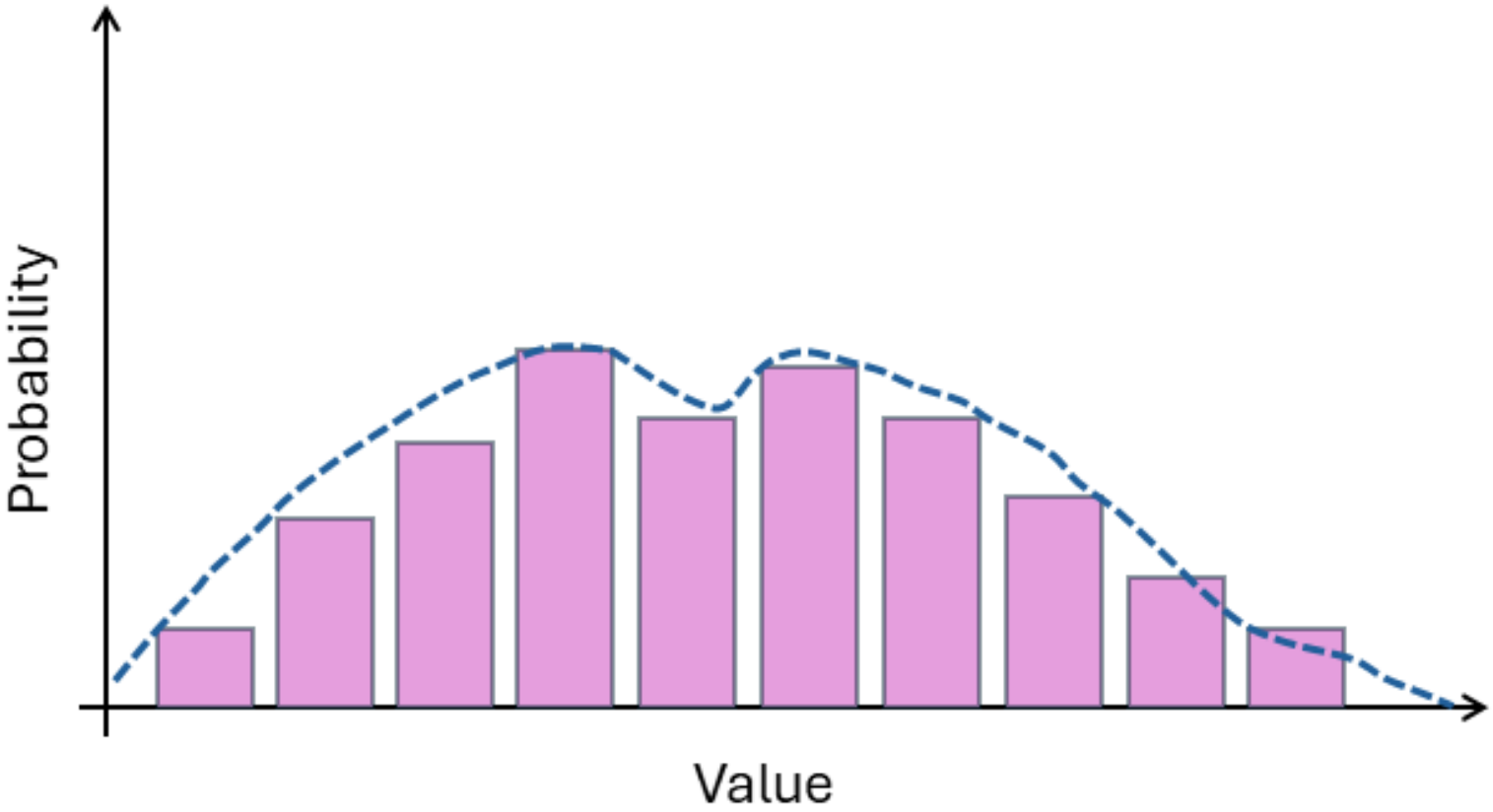


Why the Normal?



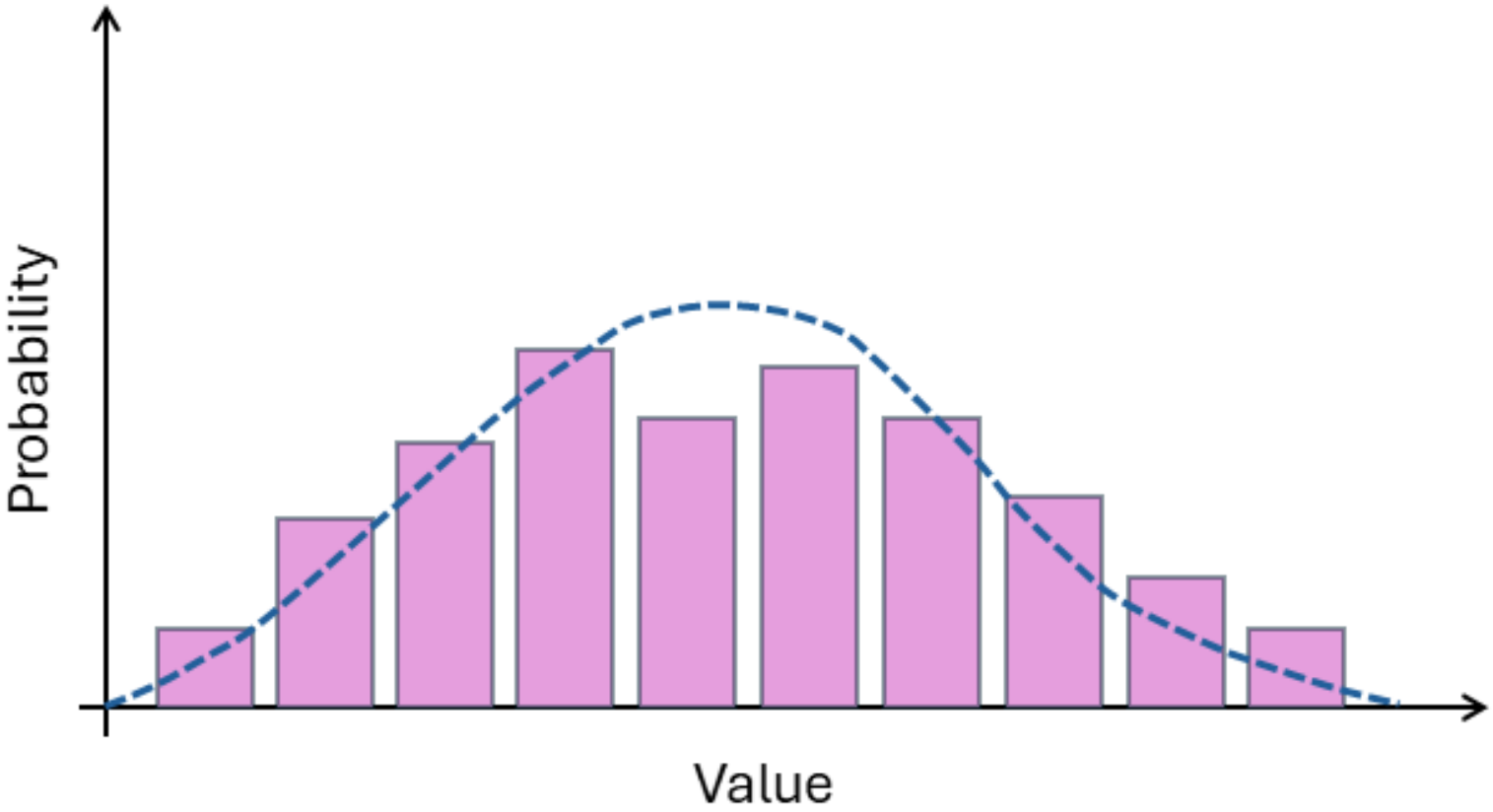
This describes the training data well, but will it generalize?

Why the Normal?



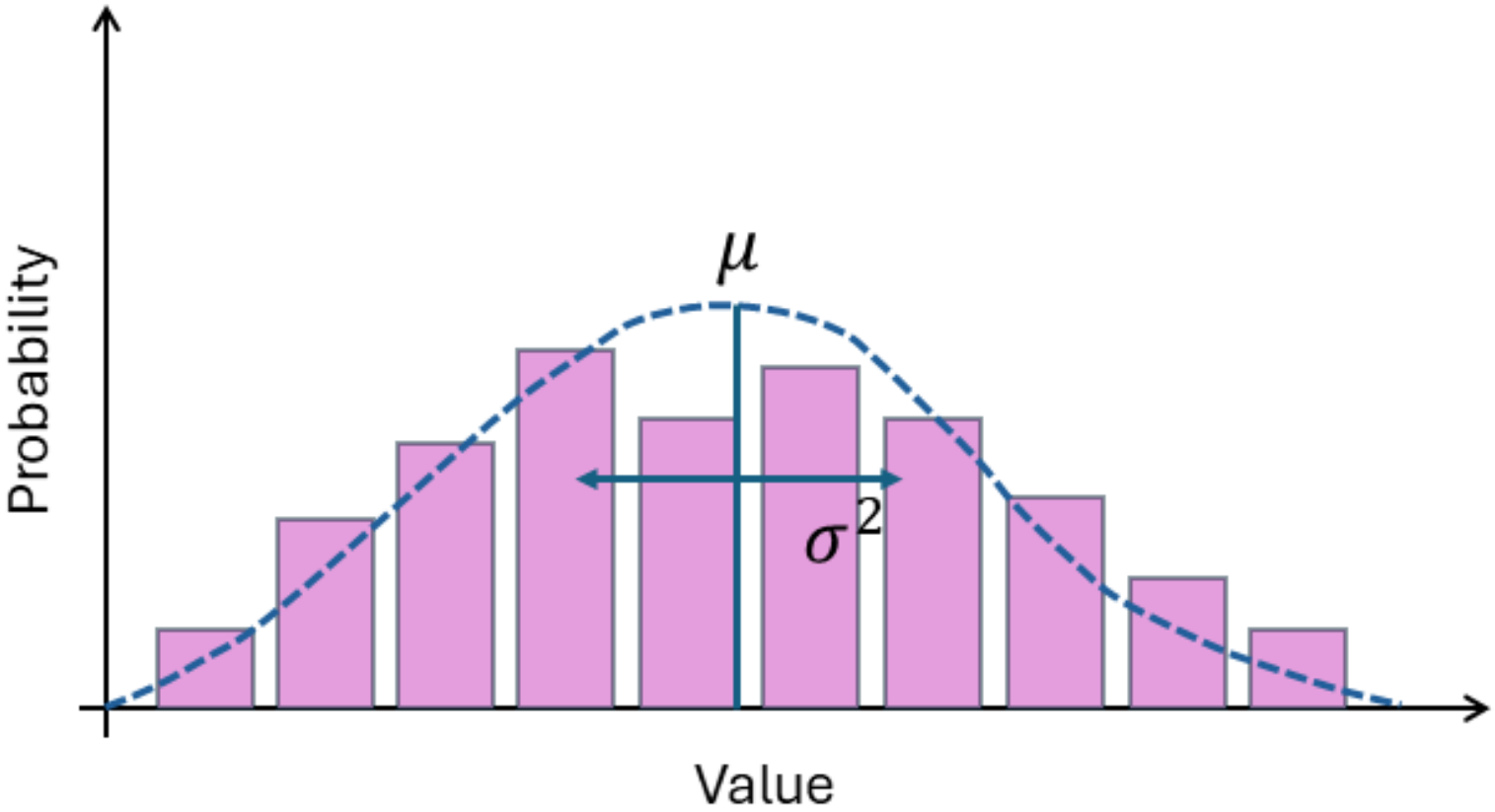
This describes the training data well, but will it generalize?
probably over fitting ...

Why the Normal?



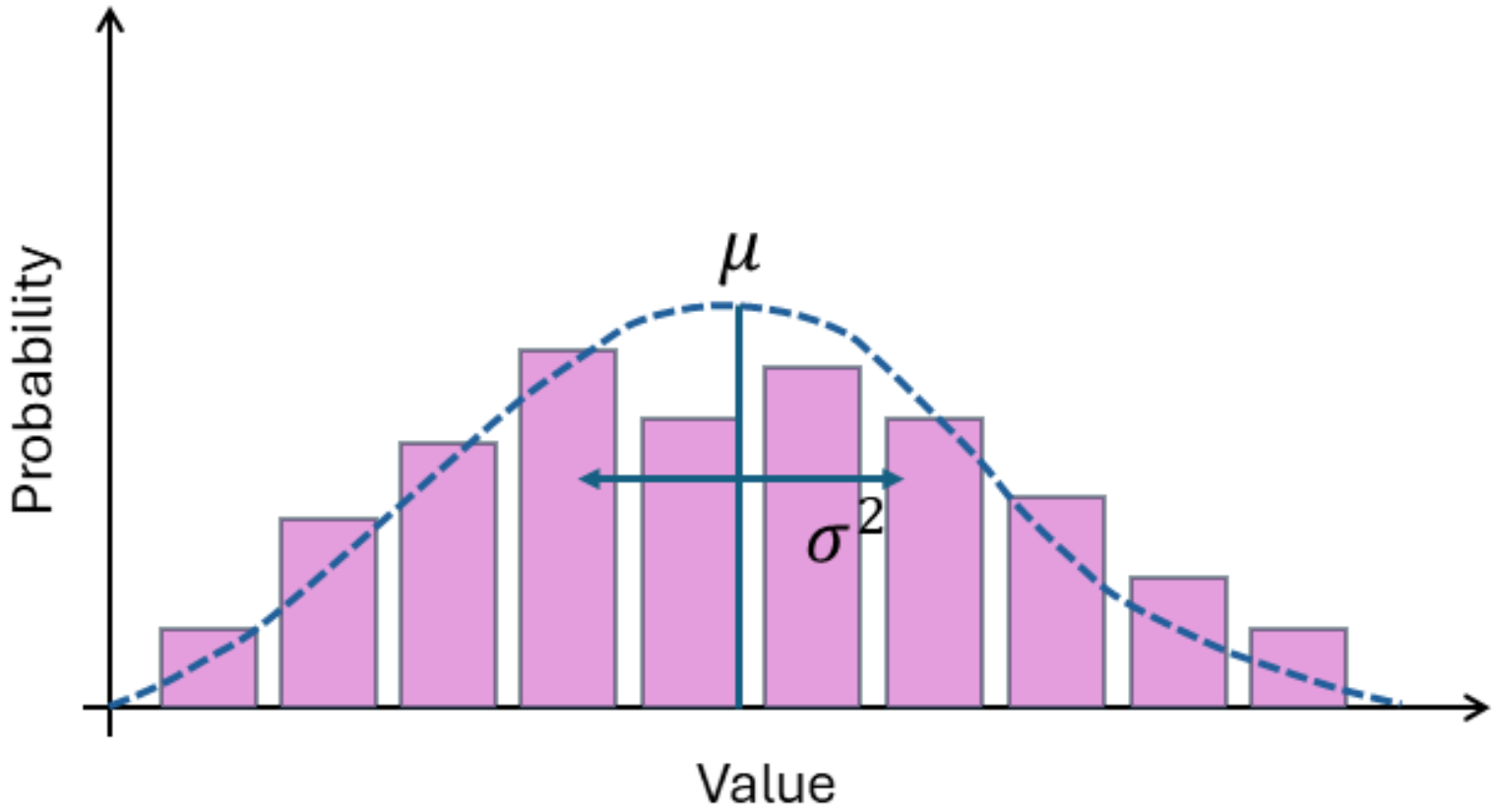
Let assume it is Gaussian!

Why the Normal?



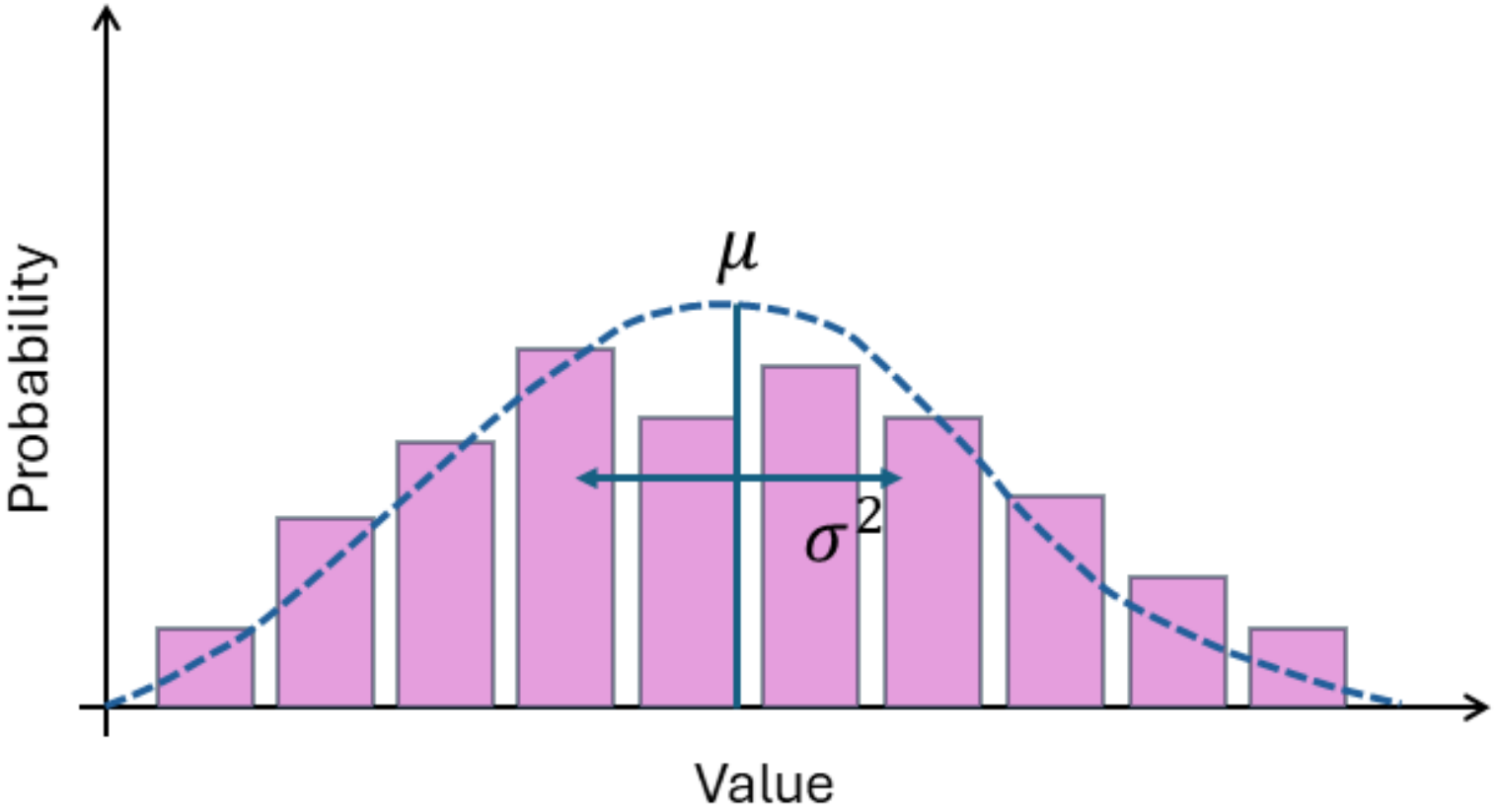
Calculate its mean and variance.

Why the Normal?



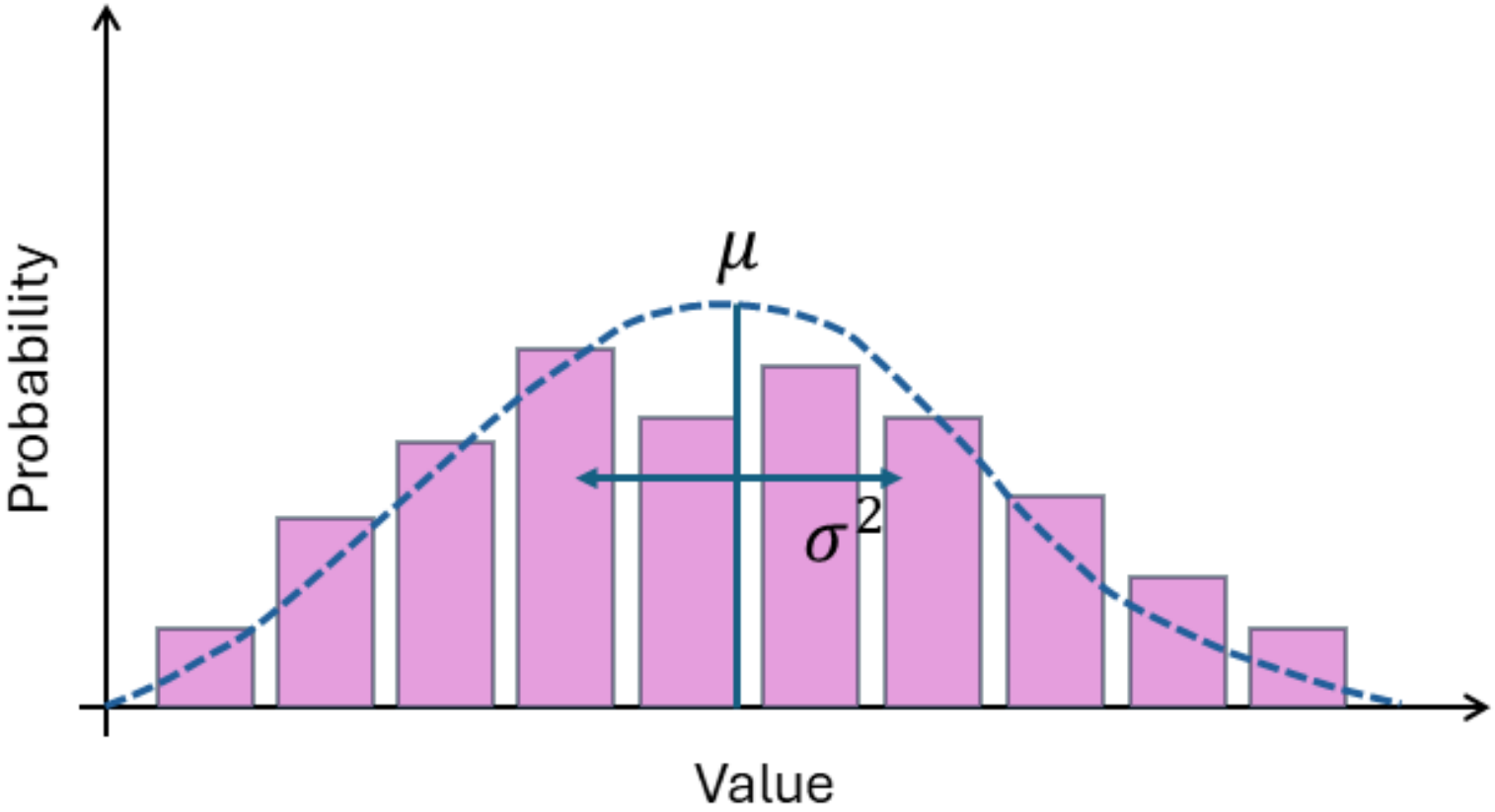
A Gaussian maximizes entropy for a given mean and variance!

Why the Normal?



Fewest assumptions. Simple. Will generalize

Why the Normal?



Fewest assumptions. Simple. Will generalize
But also, it is much easier to use ;)

Normal Probability Density Function

$$\mathcal{N}\left(\mu, \sigma^2\right)$$

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

Normal Probability Density Function

$$\mathcal{N}\left(\mu, \sigma^2\right)$$

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

Probability density at x

Normal Probability Density Function

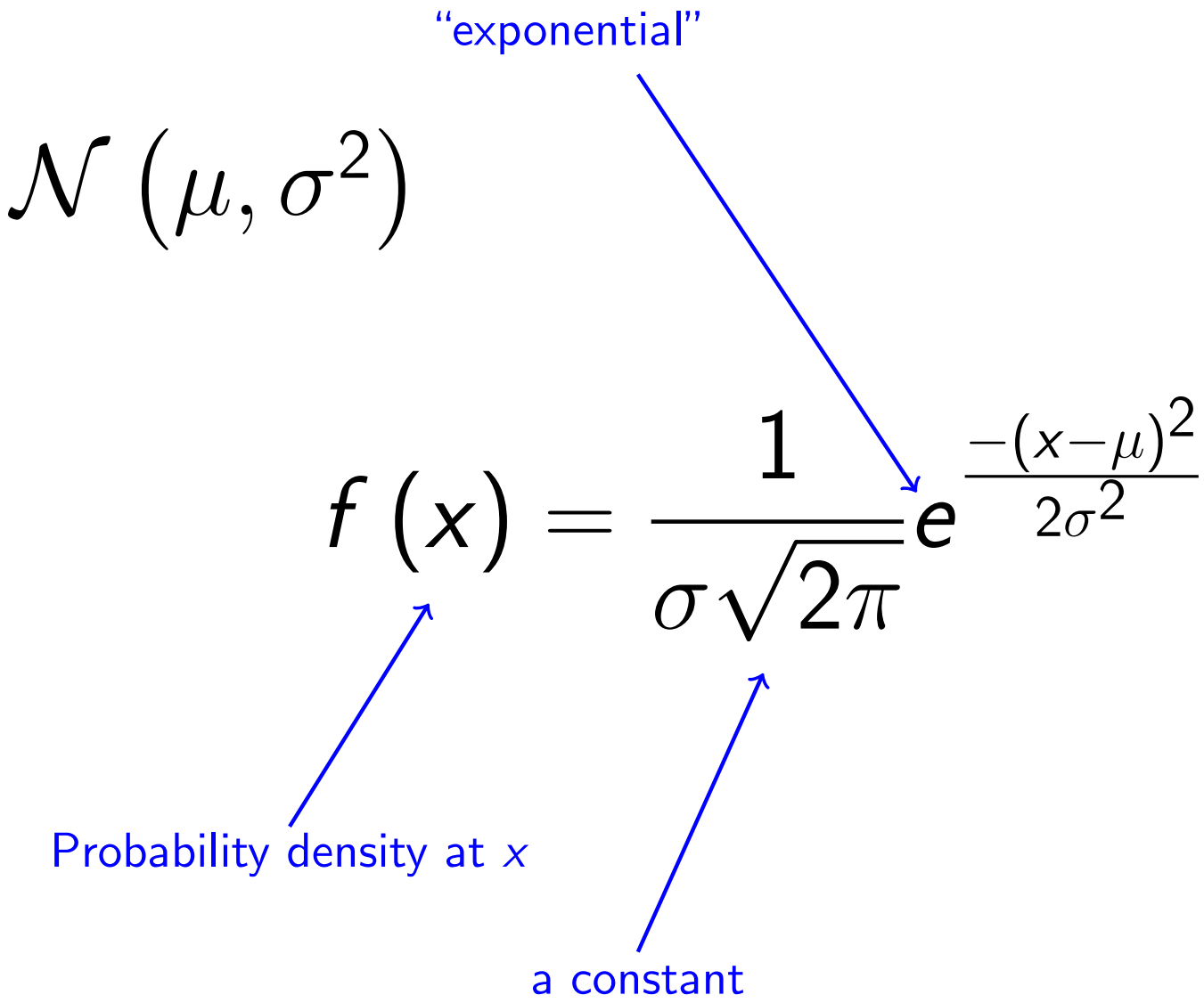
$$\mathcal{N}\left(\mu, \sigma^2\right)$$

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

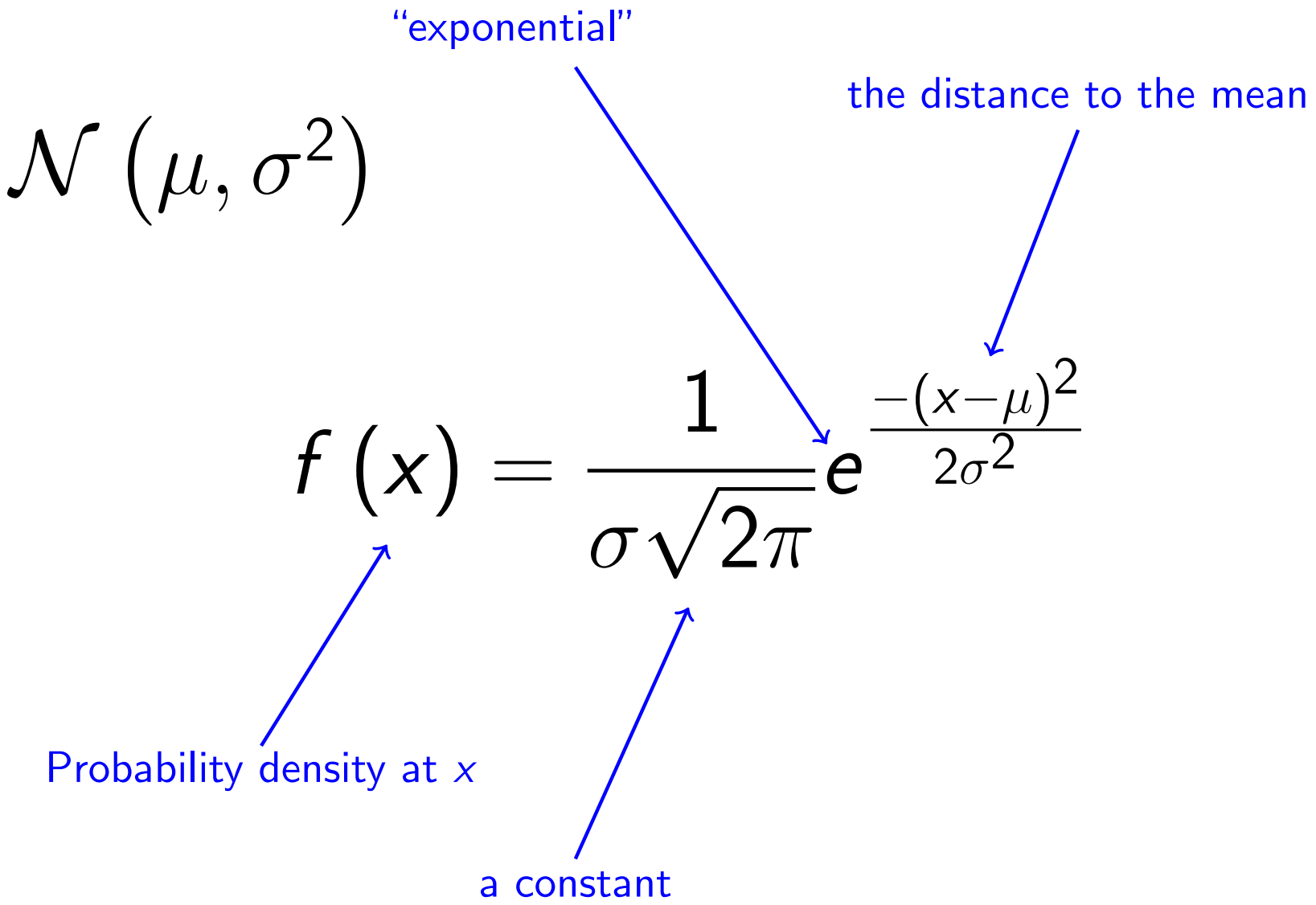
Probability density at x

a constant

Normal Probability Density Function



Normal Probability Density Function



Does it look less scary like this?

$$f(x) \propto \frac{1}{\sigma} \exp[\frac{-(x-\mu)^2}{2\sigma^2}]$$

Example

Your team is tasked with producing the side panels for Deep Sea Submarines. Physics requires all panels to be built to within 10 microns of 500. You check how precise your manufacturing is, and find these stats:

- Average panel thickness: $\mu = 500$ microns
- Variance of the thickness: $\sigma^2 = 36$ microns

What fraction of the panels you manufactured will meet standards?

Example

Your team is tasked with producing the side panels for Deep Sea Submarines. Physics requires all panels to be built to within 10 microns of 500. You check how precise your manufacturing is, and find these stats:

- Average panel thickness: $\mu = 500$ microns
- Variance of the thickness: $\sigma^2 = 36$ microns

What fraction of the panels you manufactured will meet standards?

$$X \sim \mathcal{N}\left(\mu = 500, \sigma^2 = 36\right)$$

$$P\left(490 \leq X \leq 510\right) = \int_{490}^{510} f\left(X = x\right) dx = \int_{490}^{510} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx$$

Example

No closed form for the integral!

Example

No closed form for $F(X)$!

Linear transform of Normal is Normal

Let $X \sim \mathcal{N}\left(\mu, \sigma^2\right)$

if $Y = aX + b$, then Y is also Normal

Linear transform of Normal is Normal

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

if $Y = aX + b$, then Y is also Normal

$$\begin{aligned} E[Y] &= E[aX + b] \\ &= aE[X] + b \\ &= a\mu + b \end{aligned}$$

Linear transform of Normal is Normal

$$\text{Let } X \sim \mathcal{N}\left(\mu, \sigma^2\right)$$

$$\text{if } Y = aX + b, \text{ then } Y \text{ is also Normal}$$

$$\begin{aligned} E[Y] &= E[aX + b] \\ &= aE[X] + b \\ &= a\mu + b \end{aligned}$$

$$\begin{aligned} \text{Var}(Y) &= \text{Var}(aX + b) \\ &= a^2\text{Var}(X) \\ &= a^2\sigma^2 \end{aligned}$$

Linear transform of Normal is Normal

$$\text{Let } X \sim \mathcal{N}\left(\mu, \sigma^2\right)$$

$$\text{if } Y = aX + b, \text{ then } Y \text{ is also Normal}$$

$$\begin{aligned} E[Y] &= E[aX + b] \\ &= aE[X] + b \\ &= a\mu + b \end{aligned}$$

$$\begin{aligned} \text{Var}(Y) &= \text{Var}(aX + b) \\ &= a^2\text{Var}(X) \\ &= a^2\sigma^2 \end{aligned}$$

$$Y \sim \mathcal{N}\left(a\mu + b, a^2\sigma^2\right)$$

The Cutest Linear Transform!

Let $X \sim \mathcal{N}\left(\mu, \sigma^2\right)$

There is a special case of linear transform for X:

$$Z = \frac{X - \mu}{\sigma} = \frac{1}{\sigma} X - \frac{\mu}{\sigma}$$

The Cutest Linear Transform!

Let $X \sim \mathcal{N}\left(\mu, \sigma^2\right)$

There is a special case of linear transform for X:

$$Z = \frac{X - \mu}{\sigma} = \frac{1}{\sigma}X - \frac{\mu}{\sigma} \qquad a = \frac{1}{\sigma}, b = -\frac{\mu}{\sigma}$$

The Cutest Linear Transform!

Let $X \sim \mathcal{N}\left(\mu, \sigma^2\right)$

There is a special case of linear transform for X:

$$Z = \frac{X - \mu}{\sigma} = \frac{1}{\sigma}X - \frac{\mu}{\sigma} \qquad a = \frac{1}{\sigma}, b = -\frac{\mu}{\sigma}$$
$$Z \sim \mathcal{N}\left(a\mu + b, a^2\sigma^2\right)$$

The Cutest Linear Transform!

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

There is a special case of linear transform for X:

$$Z = \frac{X - \mu}{\sigma} = \frac{1}{\sigma}X - \frac{\mu}{\sigma} \qquad a = \frac{1}{\sigma}, b = -\frac{\mu}{\sigma}$$

$$Z \sim \mathcal{N}(a\mu + b, a^2\sigma^2)$$

$$\sim \mathcal{N}\left(\frac{\mu}{\sigma} - \frac{\mu}{\sigma}, \frac{\sigma^2}{\sigma^2}\right)$$

The Cutest Linear Transform!

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

There is a special case of linear transform for X:

$$Z = \frac{X - \mu}{\sigma} = \frac{1}{\sigma}X - \frac{\mu}{\sigma} \qquad a = \frac{1}{\sigma}, b = -\frac{\mu}{\sigma}$$

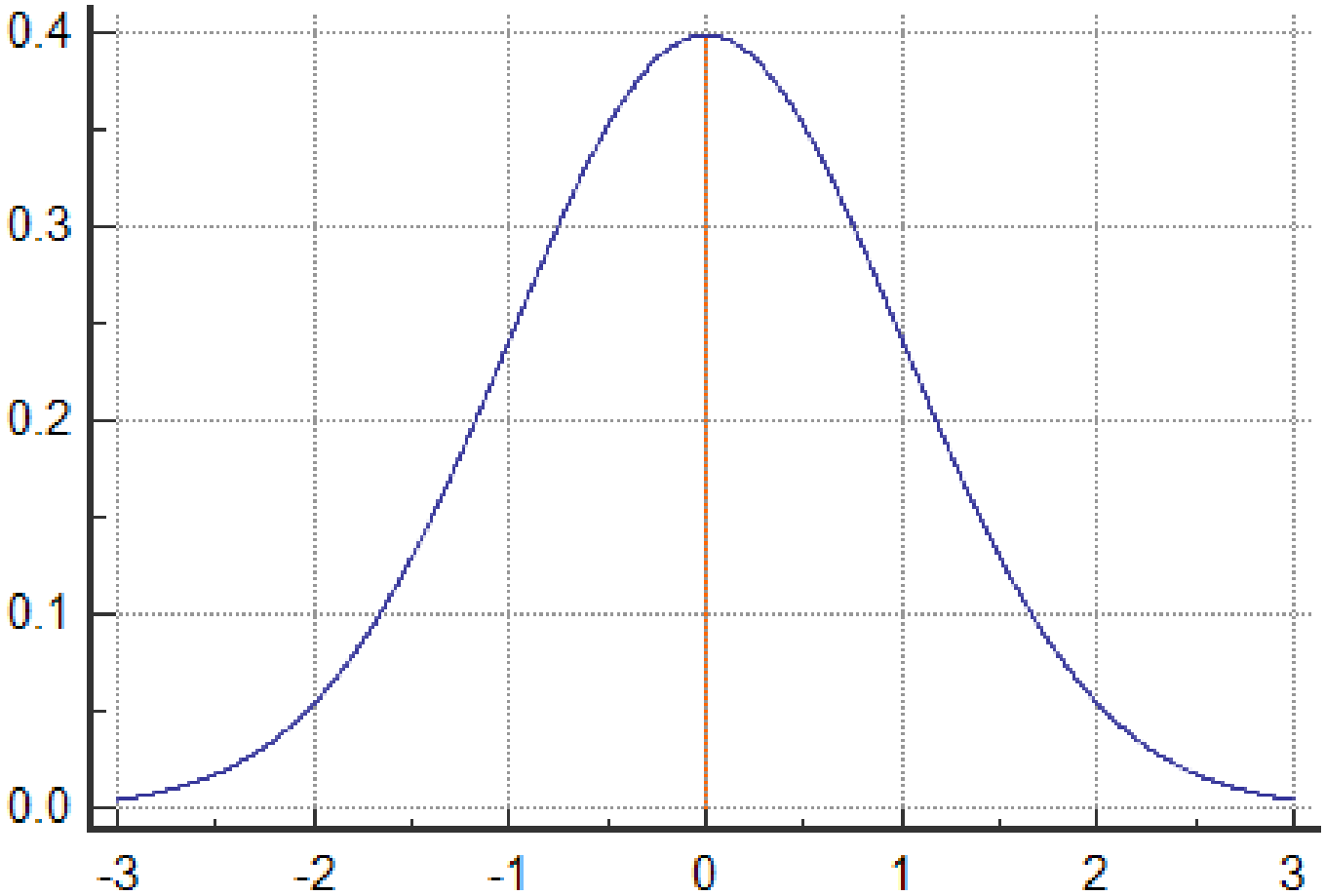
$$Z \sim \mathcal{N}(a\mu + b, a^2\sigma^2)$$

$$\sim \mathcal{N}\left(\frac{\mu}{\sigma} - \frac{\mu}{\sigma}, \frac{\sigma^2}{\sigma^2}\right)$$

$$\sim \mathcal{N}(0, 1)$$

The Standard Normal

$$Z \sim \mathcal{N}\left(\mu = 0, \sigma^2 = 1\right)$$



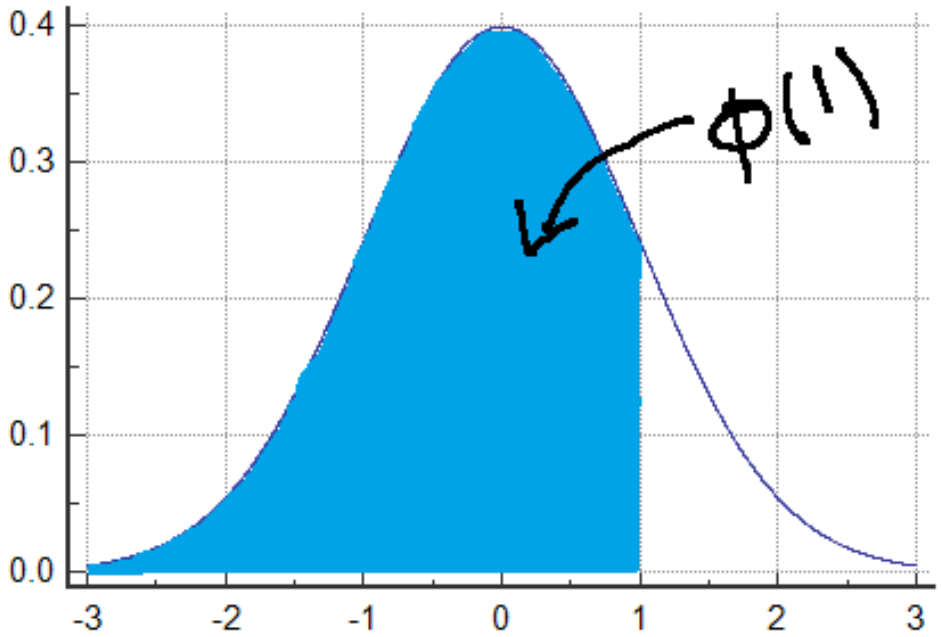
This is the PDF for the standard normal distribution

$$Z \sim \mathcal{N}\left(\mu = 0, \sigma^2 = 1\right)$$
$$F\left(x\right) = \phi\left(x\right)$$



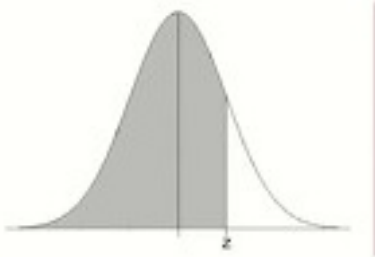
Phi

$$Z \sim \mathcal{N}\left(\mu = 0, \sigma^2 = 1\right)$$
$$F(x) = \phi(x)$$



Using Table of ϕ

Standard Normal Cumulative Probability Table

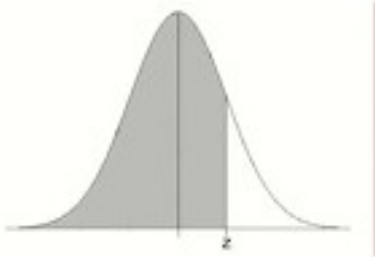


Cumulative probabilities for POSITIVE z-values are shown in the following table:

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319

Using Table of ϕ

Standard Normal Cumulative Probability Table

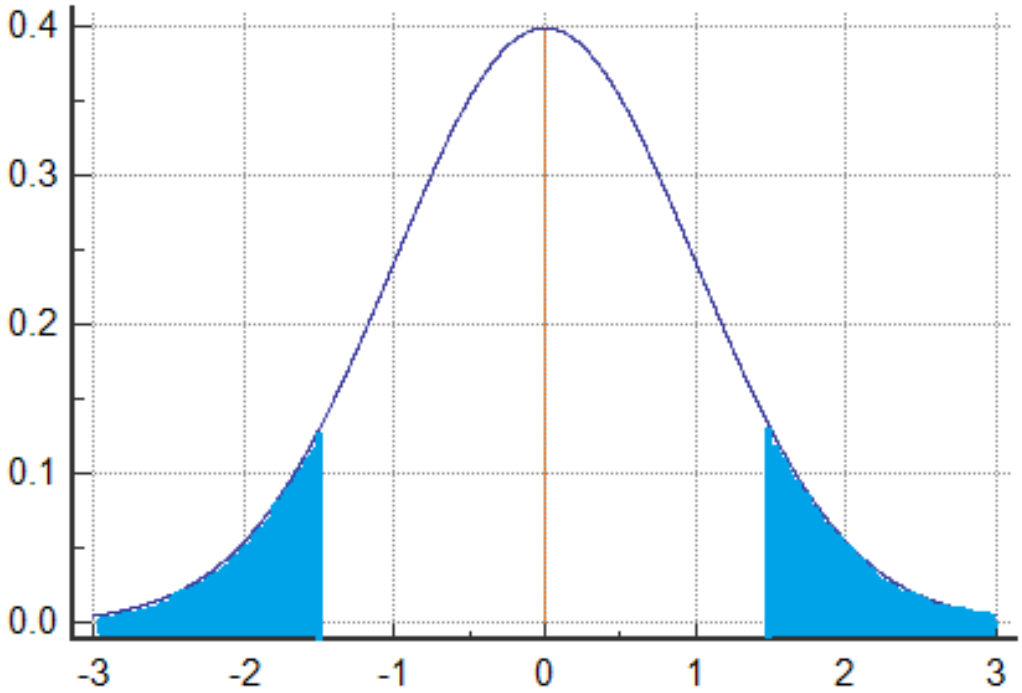


Cumulative probabilities for POSITIVE z-values are shown in the following table:

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319

Symmetry of *phi*

$$\phi(-a) = 1 - \phi(a)$$



Interval of ϕ

$$P(c < Z < d) = \phi(d) - \phi(c)$$

Compute $F(X)$ via Transform

Let $X \sim \mathcal{N}(\mu, \sigma^2)$ $\qquad$ $Z = \frac{X-\mu}{\sigma}$

Use Z to compute $F(X)$

Compute $F(X)$ via Transform

Let $X \sim \mathcal{N}(\mu, \sigma^2)$ $\qquad Z = \frac{X-\mu}{\sigma}$

Use Z to compute $F(X)$

$$F_X(x) = P(X \leq x)$$

Compute $F(X)$ via Transform

Let $X \sim \mathcal{N}(\mu, \sigma^2)$ $\qquad Z = \frac{X-\mu}{\sigma}$

Use Z to compute $F(X)$

$$F_X(x) = P(X \leq x)$$
$$= P(X - \mu \leq x - \mu)$$

Compute $F(X)$ via Transform

Let $X \sim \mathcal{N}(\mu, \sigma^2)$ $\qquad Z = \frac{X-\mu}{\sigma}$

Use Z to compute $F(X)$

$$
\begin{aligned}
F_X(x) &= P(X \leq x) \\
&= P(X - \mu \leq x - \mu) \\
&= P\left(\frac{X-\mu}{\sigma} \leq \frac{x-\mu}{\sigma}\right)
\end{aligned}
$$

Compute $F(X)$ via Transform

Let $X \sim \mathcal{N}(\mu, \sigma^2)$ $\qquad Z = \frac{X-\mu}{\sigma}$

Use Z to compute $F(X)$

$$
\begin{aligned}
F_X(x) &= P(X \leq x) \\
&= P(X - \mu \leq x - \mu) \\
&= P\left(\frac{X-\mu}{\sigma} \leq \frac{x-\mu}{\sigma}\right) \\
&= P\left(Z \leq \frac{x-\mu}{\sigma}\right)
\end{aligned}
$$

Compute $F(X)$ via Transform

Let $X \sim \mathcal{N}(\mu, \sigma^2)$ $\qquad Z = \frac{X-\mu}{\sigma}$

Use Z to compute $F(X)$

$$
\begin{aligned}
F_X(x) &= P(X \leq x) \\
&= P(X - \mu \leq x - \mu) \\
&= P\left(\frac{X-\mu}{\sigma} \leq \frac{x-\mu}{\sigma}\right) \\
&= P\left(Z \leq \frac{x-\mu}{\sigma}\right) \\
&= \phi\left(\frac{x-\mu}{\sigma}\right)
\end{aligned}
$$

Takeaway Message

For normal distribution, $F(x)$ is computed using the ϕ transform

And Here We Are!

$$\mathcal{N}\left(\mu, \sigma^2\right)$$

CDF of Standard Normal: A function that has been solved for numerically

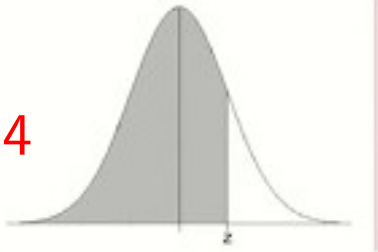
$$F(x) = \Phi\left(\frac{x-\mu}{\sigma}\right)$$

The cumulative density function of any normal

Using the Phi Table

Standard Normal Cumulative Probability Table

$$\phi(0.54) = 0.7054$$



Cumulative probabilities for POSITIVE z-values are shown in the following table:

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319

Do We Have To Use The Table?

Standard Normal Cumulative Probability Table

Cumulative probabilities for POSITIVE z-values are shown in the following table:

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319

No, it's kinda old school!

Do We Have To Use The Table?

The scipy library in Python

stats.norm.cdf(x, mean, std)

$$= P(X < x) \text{ where } X \sim \mathcal{N}(\mu, \sigma^2)$$
$$= \phi\left(\frac{x-\mu}{\sigma}\right)$$