INNOMATICS TECHNOLOGY HUB

A sister concern of

KenMinds
...committed to solution excellence

Review

- Data
  - Type of Data
  - Central Tendencies, Spread, Data Presentation – Box Plot
- Probability
  - Probability Rules, Venn Diagram
  - Conditional Probability
  - Bayes Theorem
  - Confusion Matrix & Classification Errors
- Probability Distributions
  - Connection to Histogram

Probability Review

- A pair of dice are rolled. What is the probability that the sum of the dice is 6 ?

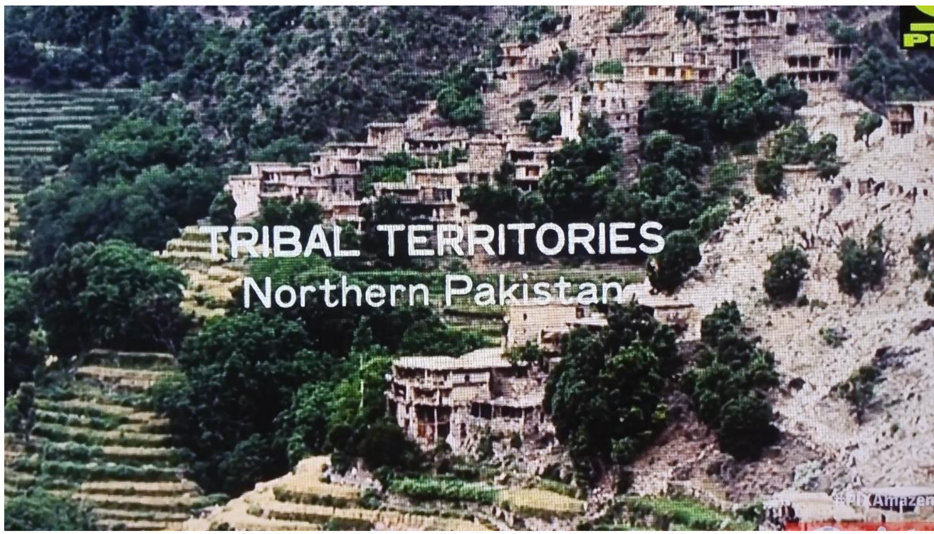
  - Total number of possibilities = 36
  - Sum 7 events =

  (1,5), (2,6) , (3,3), (4,2), (5,1)

  - Probability = 5 / 36

Bayes’ Theorem: Finding Osama Bin Laden

“Bin Laden was equality likely to be in a compound in a city or in the mountain caves of tribal areas”



Bayes’ Theorem: Finding Osama Bin Laden

The criteria to mark a compound as positive include high walls, inmates excepting a messenger not entering or leaving the compound, inmates carefully avoiding detection including by surveillance.

Given a compound meets all the criteria what is the probability that Bin Laden is in that compound ?

Find Osama Bin Laden

P(Bin Laden is in Compound) = 0.5

P(Compound is +ve | Bin Laden is in Compound) = 0.9

Probability that the security is very high given that Bin Laden is in the Compound = 0.9

P(Compound is +ve | Bin Laden is NOT in Compound) = 0.6

Probability that the security is very high given that Bin Laden is in the Compound = 0.9

Find Osama Bin Laden

$$\frac{P\ Bin\ Laden\ is\ in\ Compound\ *\ P(Compound\ is\ +|Bin\ Laden\ is\ in\ Compound)}{(P\ (Bin\ Laden\ is\ in\ Compound)*\ P(Compound\ is\ +\ \ |\ Bin\ Laden\ is\ in\ Compound\ \ +) \\ + \\ P(Bin\ Laden\ is\ NOT\ in\ Compound)*\ P(Compound\ is\ +\ |Bin\ Laden\ is\ NOT\ in\ Compound)}$$

P(Bin Laden is in Compound | Compound is +ve)

= $\frac{05\ *0.9}{0.5\ *0.9+0.5\ *0.6} = 60\ \%$

Find Osama Bin Laden

P(Bin Laden is in Compound) = 0.5

P(Compound is +ve | Bin Laden is in Compound) = 0.9

P(Compound is +ve | Bin Laden is NOT in Compound) = 0.05

P(Bin Laden is in Compound | Compound is +ve)

$$= \frac{05 * 0.9}{0.5 * 0.9 + 0.5 * 0.05} = 95\ \%$$

Distributions

Geometric: For estimating number of attempts before first success.

Binomial: For estimating number of success in ***n*** attempts

Poisson: For estimating ***n*** number of events in a given time period when on average we see ***m*** events.

Exponential: Time between events

Probability Distribution

Here are a few scenarios. Identify the distribution and calculate expectation, variance and the required probabilities.

Q1. In an Archery competition a man has 0.3 probability of hitting bulls-eye. If he has 10 tries, what is the probability he will hit bulls-eye less than 3 times?

Q2. On average, 2 buses stops at a certain point in any given 15 minute period. What is the probability that no buses will turn up in a single 15 minute interval?

Q3. 20% of cereal packets contain a free toy. What is the probability you will need to open fewer than 4 cereal packets before finding your first toy?

Probability Distribution

Q1. In an Archery competition a man has 0.3 probability of hitting bulls-eye. If he has 10 tries, what is the probability he will hit bulls-eye less than 3 times?

Solution:

p = 0.3 , q = 1-p = 0.7, n = 10, r = 3

P(x < 3) = P(x = 0) + P(x = 1) + P(x = 2)

P(x = r) = ${}^{n}_{r}c * \ p^{r} * q^{n-r}$

P(x <3) = 0.382

Probability Distribution

Q2. On average, 2 buses stops at a certain point in any given 15 minute period. What is the probability that no buses will turn up in a single 15 minute interval?

Solution:

E(x) = $\lambda = 2$

P(x=r) = $\frac{\lambda^r}{r!} e^{-\lambda}$

P(x = 0) = $\frac{2^0}{0!} e^{-2}$

= e^{-2}

Probability Distribution

Q3. 20% of cereal packets contain a free toy. What is the probability you will need to open fewer than 4 cereal packets before finding your first toy?

Solution:

$p = 0.2$, $q = 1- 0.2 = 0.8$, $r < 4$ or ≤ 3

$E(x) = \frac{1}{p} = 5$

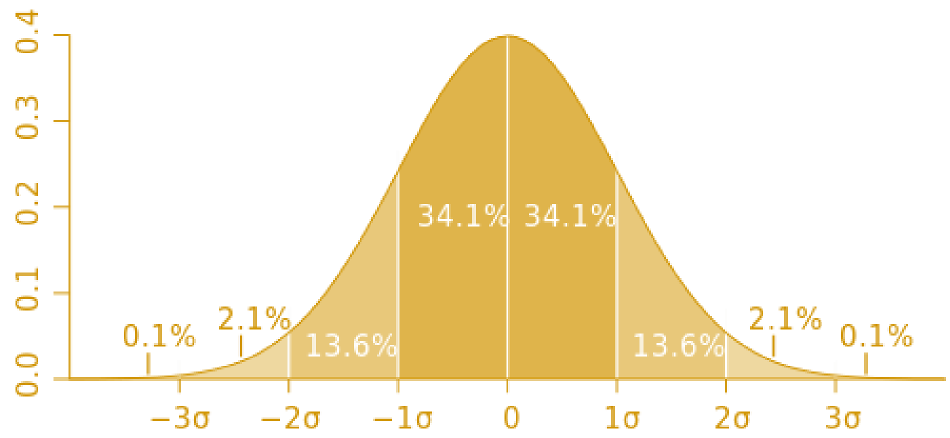
$Var(x) = \frac{q}{p^2} = 20$

$P(x \leq r) = 1 - q^r$

$P(x \leq r) = 0.488$

https://www.mathsisfun.com/data/quincunx.html

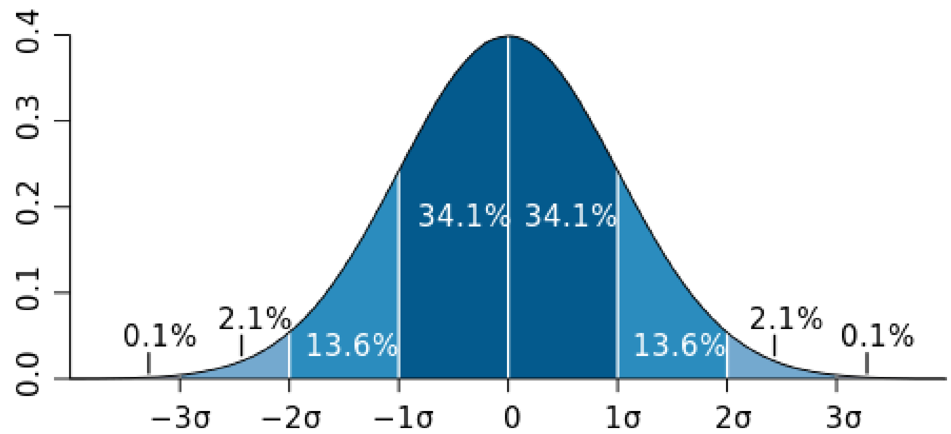
https://www.mathsisfun.com/data/standard-normal-distribution-table.html



Using the Normal Distribution

Normal (Gaussian) Distribution

- Mean = Median = Mode
- Zero Skew and Kurtosis = 3
- X ~ N(μ, σ^2)
- 68.95 – 99.7 empirical rule
- Shaded area gives the probability that X is between the corresponding values

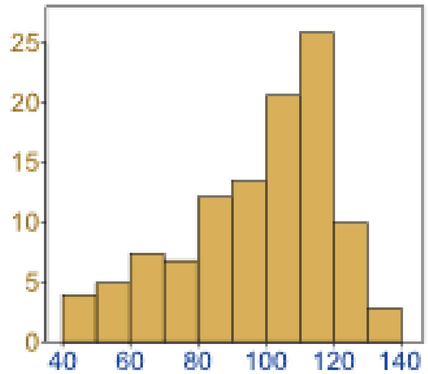


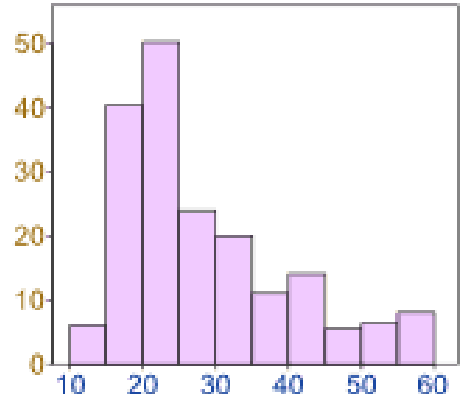
$$f(x, \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}}\, e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

Measure of Spread

Data can be "distributed" (spread out) into different ways.

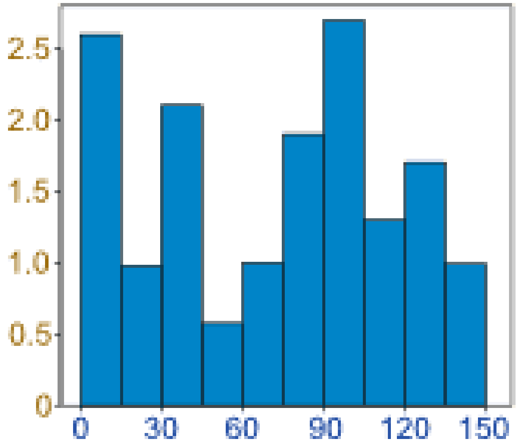
It can be spread out more on the **left**





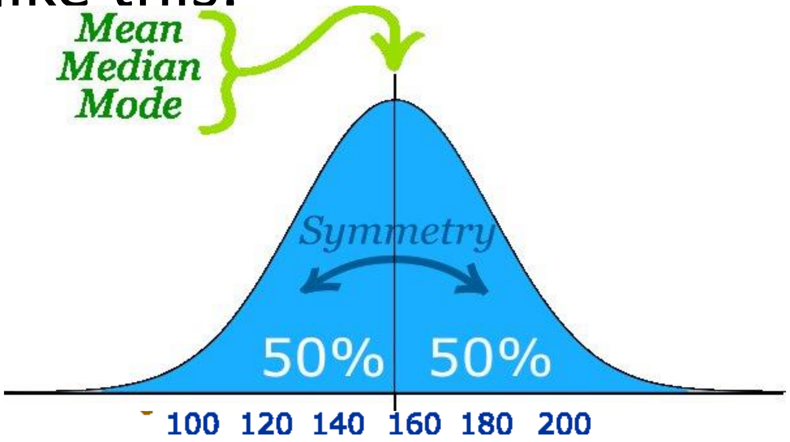
It can be spread out more on the **right**

Or it can be **jumbled** up



Measure of Spread

But there are many cases where the data tends to be around a central value with no bias left or right, and it gets close to a "**Normal Distribution**" like this:

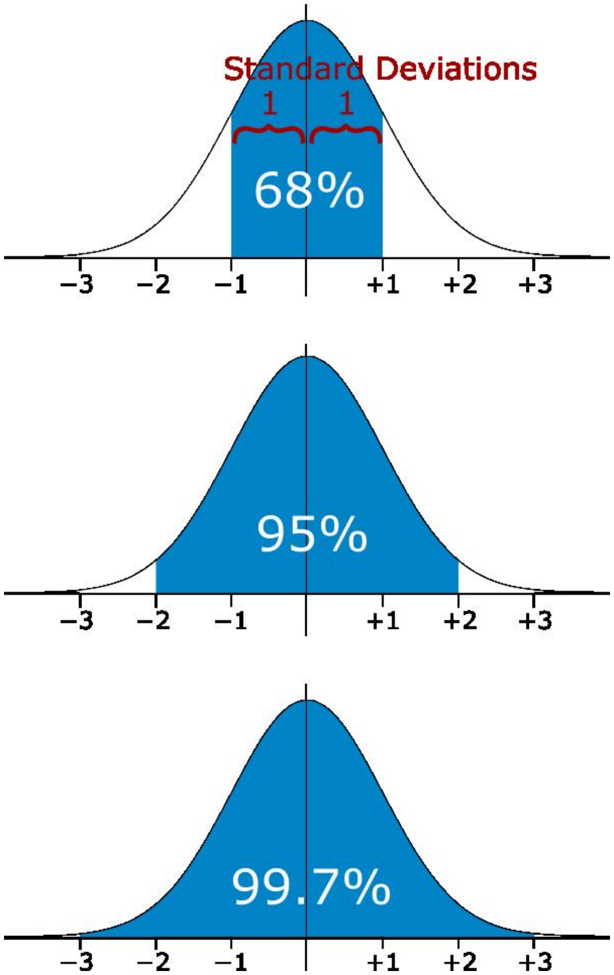


Normal Distribution

Many things closely follow a Normal Distribution:

- heights of people
- size of things produced by machines
- errors in measurements
- blood pressure
- marks on a test

Measure of Spread

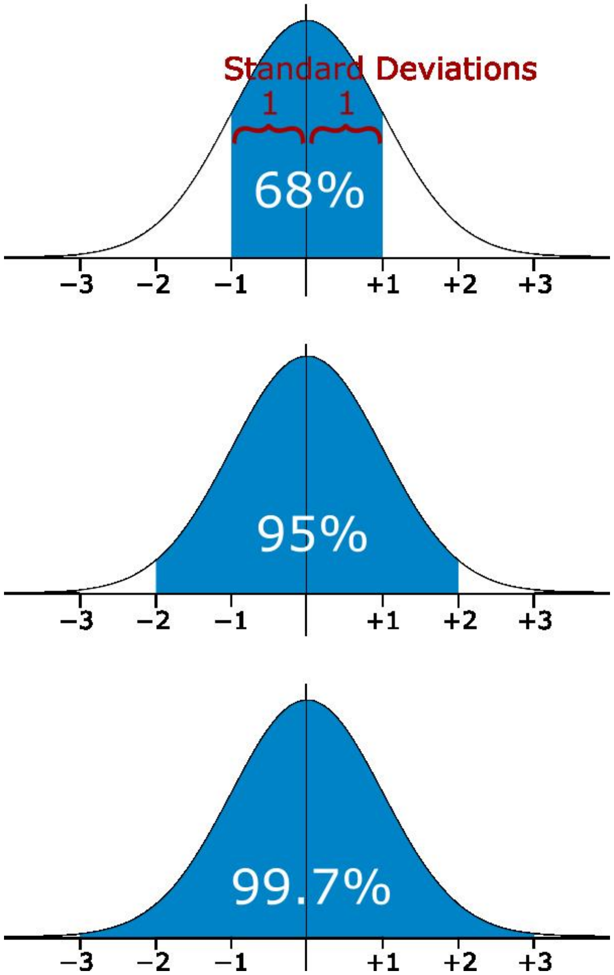


Standard Deviation

The **Standard Deviation** is a measure of how spread out numbers are :

i.e., *it is the average spread of data around it mean*

Measure of Spread



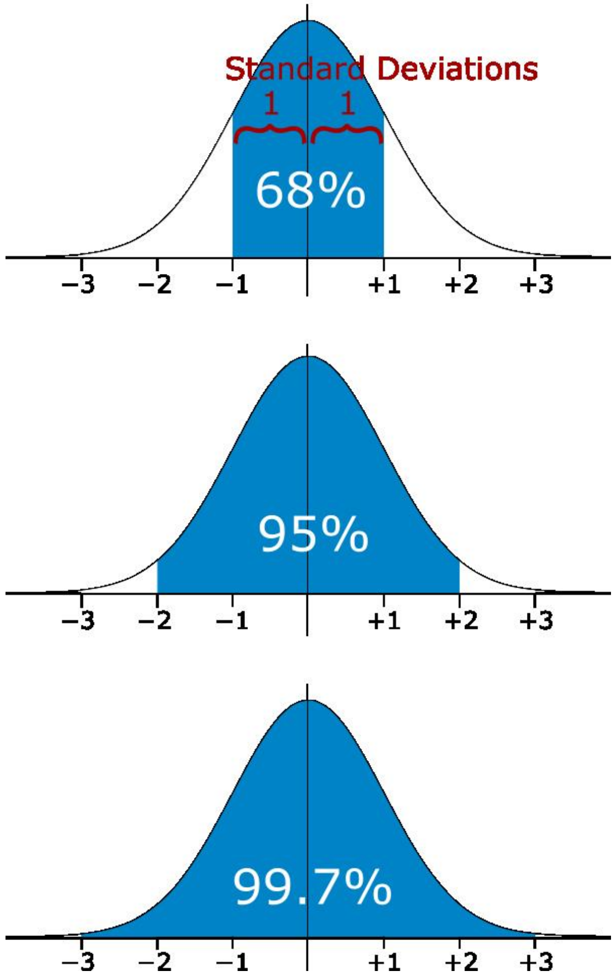
Standard Deviation

68% of values are within
1 standard deviation of the mean

95% of values are within
2 standard deviations of the mean

99.7% of values are within
3 standard deviations of the mean

Measure of Spread



You know 68 – 95 – 99 rule.

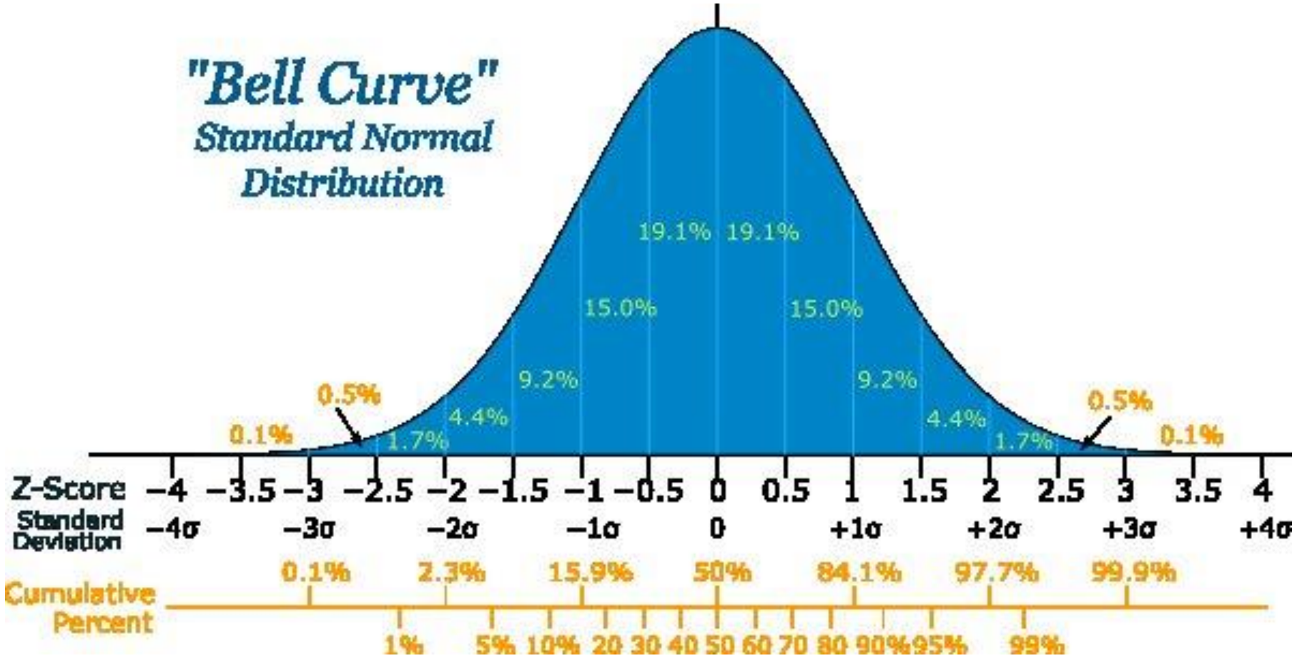
A company produces a lightweight values that is specified to weigh **1500g,** but there are imperfections in the process. While the mean weight is 1500g, the standard deviation is 300g.

Q1. What is the range of weights within which 95% of the values will fall ?

Q2. Approximately 16% of the weights will be more than what value ?

Q3. Approximately 0.15% of the weight will be less than what value ?

Measure of Spread

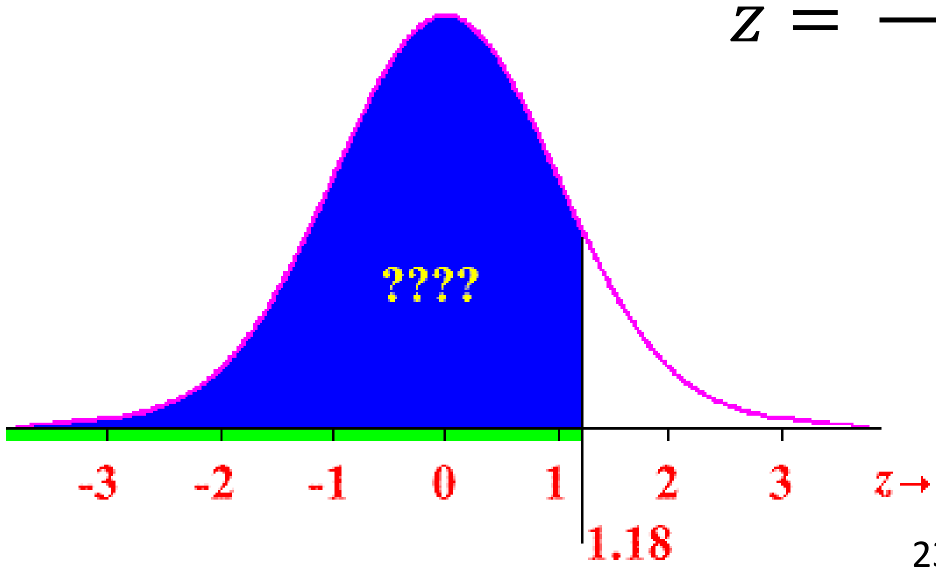


Standardize a Normal Distribution

No closed form formula exits for area under curve for the Gaussian. Hence we need to use pre-computed tables for obtaining area under the curve.

We rescale the Gaussian distribution using the following transformation.

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

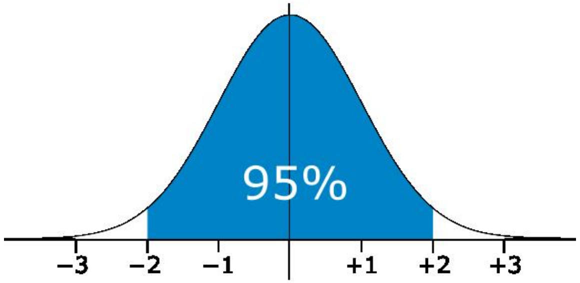


Example:

95 % of students at a school are between 1.1 m and 1.7 m tall. Assuming this data is normally distributed, calculate the mean and standard deviation.

Example:

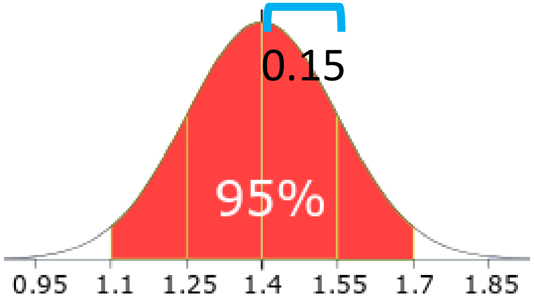
95 % of students at a school are between 1.1 m and 1.7 m tall. Assuming this data is normally distributed, calculate the mean and standard deviation.



The mean is halfway between 1.1m and 1.7m:
Mean = (1.1m + 1.7m) / 2 = **1.4m**

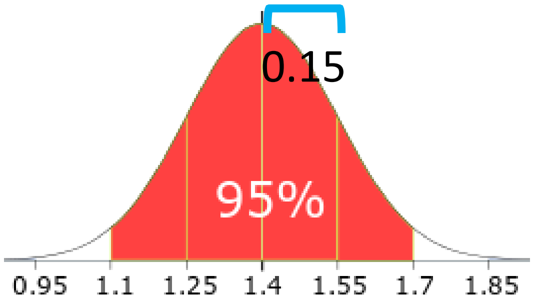
95% is 2 standard deviations either side of the mean (a total of 4 standard deviations) so:

1 standard deviation = (1.7m-1.1m) / 4
= 0.6m / 4
= **0.15m**



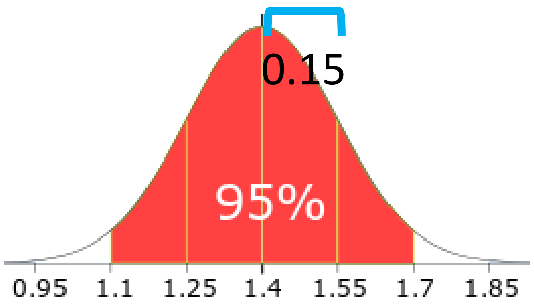
Example:

In that same school one of you friends is 1.85m tall. What is the Z-score of your friend's height ?



Example:

In that same school one of you friends is 1.85m tall. What is the Z-score of your friend's height ? (*How far is 1.85 from the mean?)*



It is 1.85 - 1.4 = **0.45m from the mean**

How many standard deviations is that? The standard deviation is 0.15m, so:

0.45m / 0.15m = **3 standard deviations**

Gaussian For Blind date

Julie is a student, and her best friend keeps trying to get her fixed up on blind dates in the hope that she'll find that special someone.

- She want to date only tall guys

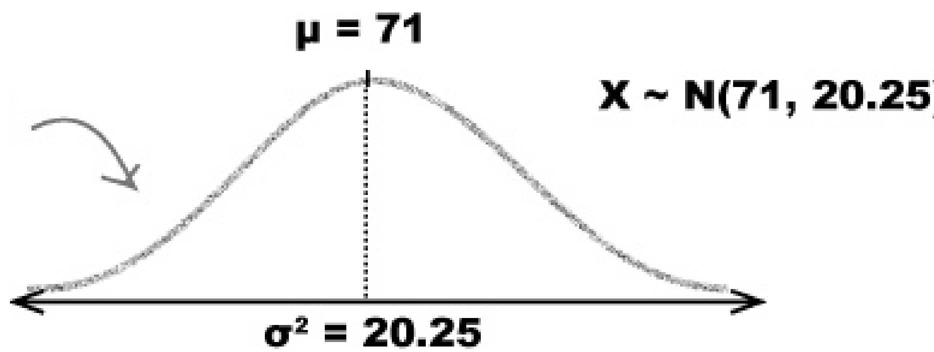
Oh! By the way Julie is 64" tall.

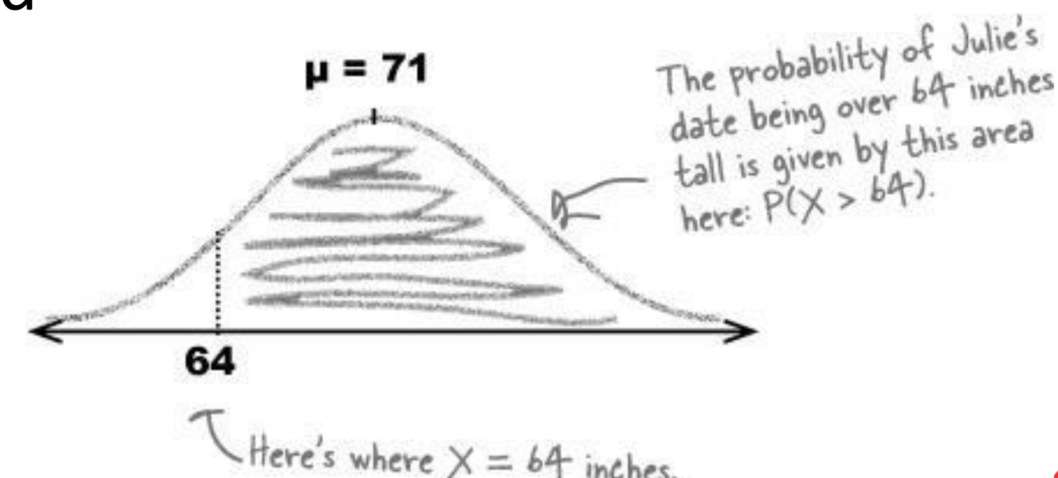
Probability Distribution

- Step 1: Determine the distribution

Julie wants to marry a person taller than her and is going on blind dates. The mean height of the 'available' guys is 71" and the variance is 20.25 $inch^2$.

Julie is 64" tall.

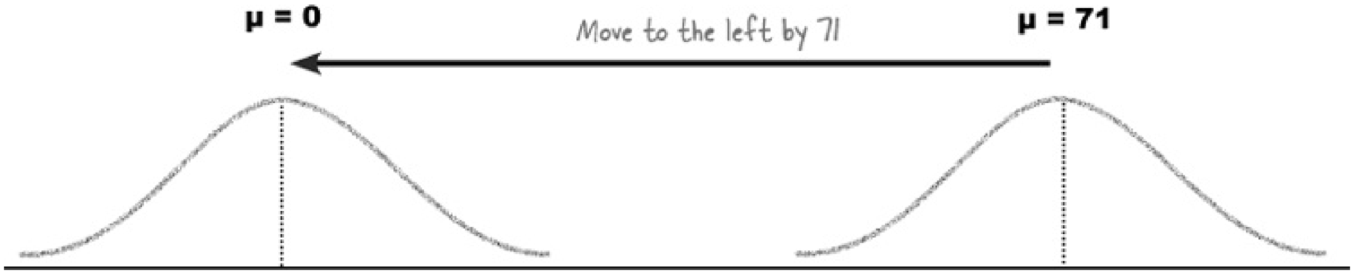




Calculating Normal Probabilities

- Step 2: Standardize to Z ~ N(0,1)

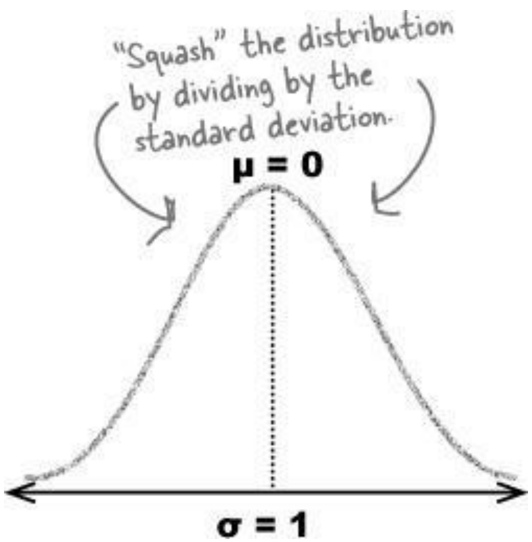
1. Move the mean This gives a new distribution X-71 ~ N(0,20.25)



2. Rescale the width by dividing by the standard deviation

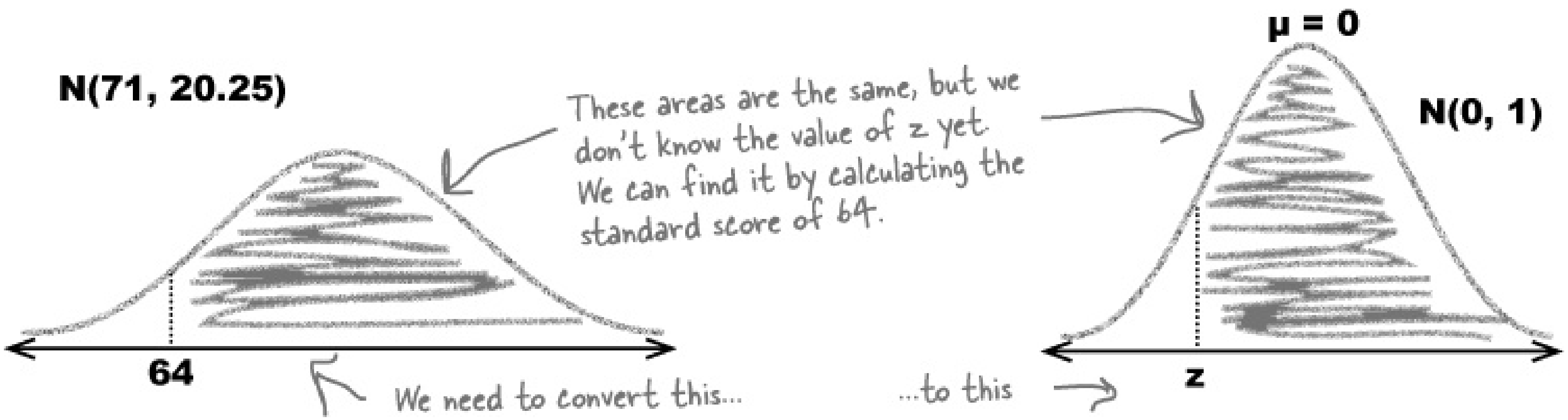
This gives us $\frac{X-71}{4.5}$

$Z = \frac{X-\mu}{\sigma}$ is called standard score or the Z-score



Calculating Normal Probabilities

- Step 2: Standardize to Z ~ N(0,1)



$$Z = \frac{X - 71}{4.5} = -\ 1.56 \text{ for height } 64''$$

Calculating Normal Probabilities

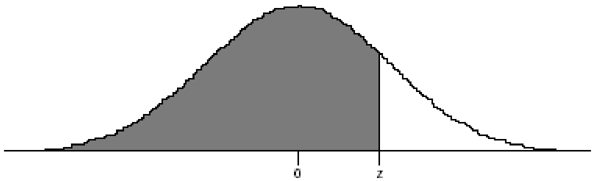
- Step 3: Look up the
- probability in the tables

Note the tables give P(Z<z)

$$Z = \frac{X - 71}{4.5} = -1.56$$

for height 64"

$$P(Z > -1.56) = 1 - P(Z < -1.56)$$
$$= 1 - 0.0594$$
$$= 0.9406$$



Normal Deviate z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
-4.0	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
-3.9	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
-3.8	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
-3.7	.0001	.0001	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
-3.6	.0002	.0002	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001
-3.5	.0002	.0002	.0002	.0002	.0002	.0002	.0002	.0002	.0002	.0002
-3.4	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0002
-3.3	.0005	.0005	.0005	.0004	.0004	.0004	.0004	.0004	.0004	.0003
-3.2	.0007	.0007	.0006	.0006	.0006	.0006	.0006	.0005	.0005	.0005
-3.1	.0010	.0009	.0009	.0009	.0008	.0008	.0008	.0008	.0007	.0007
-3.0	.0013	.0013	.0013	.0012	.0012	.0011	.0011	.0011	.0010	.0010
-2.9	.0019	.0018	.0018	.0017	.0016	.0016	.0015	.0015	.0014	.0014
-2.8	.0026	.0025	.0024	.0023	.0023	.0022	.0021	.0021	.0020	.0019
-2.7	.0035	.0034	.0033	.0032	.0031	.0030	.0029	.0028	.0027	.0026
-2.6	.0047	.0045	.0044	.0043	.0041	.0040	.0039	.0038	.0037	.0036
-2.5	.0062	.0060	.0059	.0057	.0055	.0054	.0052	.0051	.0049	.0048
-2.4	.0082	.0080	.0078	.0075	.0073	.0071	.0069	.0068	.0066	.0064
-2.3	.0107	.0104	.0102	.0099	.0096	.0094	.0091	.0089	.0087	.0084
-2.2	.0139	.0136	.0132	.0129	.0125	.0122	.0119	.0116	.0113	.0110
-2.1	.0179	.0174	.0170	.0166	.0162	.0158	.0154	.0150	.0146	.0143
-2.0	.0228	.0222	.0217	.0212	.0207	.0202	.0197	.0192	.0188	.0183
-1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0239	.0233
-1.8	.0359	.0351	.0344	.0336	.0329	.0322	.0314	.0307	.0301	.0294
-1.7	.0446	.0436	.0427	.0418	.0409	.0401	.0392	.0384	.0375	.0367
-1.6	.0548	.0537	.0526	.0516	.0505	.0495	.0485	.0475	.0465	.0455
-1.5	.0668	.0655	.0643	.0630	.0618	.0606	.0594	.0582	.0571	.0559
-1.4	.0808	.0793	.0778	.0764	.0749	.0735	.0721	.0708	.0694	.0681
-1.3	.0968	.0951	.0934	.0918	.0901	.0885	.0869	.0853	.0838	.0823
-1.2	.1151	.1131	.1112	.1093	.1075	.1056	.1038	.1020	.1003	.0985
-1.1	.1357	.1335	.1314	.1292	.1271	.1251	.1230	.1210	.1190	.1170
-1.0	.1587	.1562	.1539	.1515	.1492	.1469	.1446	.1423	.1401	.1379

Z-table

Attention Check

Find the probability **Julie** finds a man in the between 66" to 76".

Z-Score @ 66" = (66 – 71)/4.5 = - 1.11

P(Z = -1.11) = 0.1131

Z-Score @ 76" = (76 - 71)/4.5 = 1.11

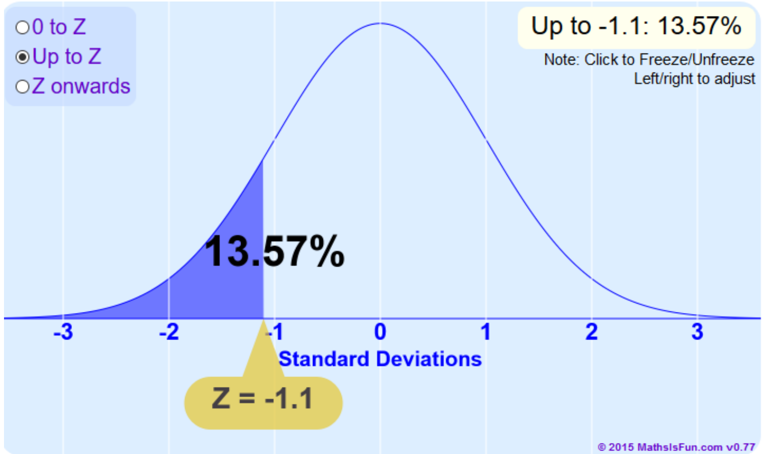
P(Z = 1.11) = 0.8665

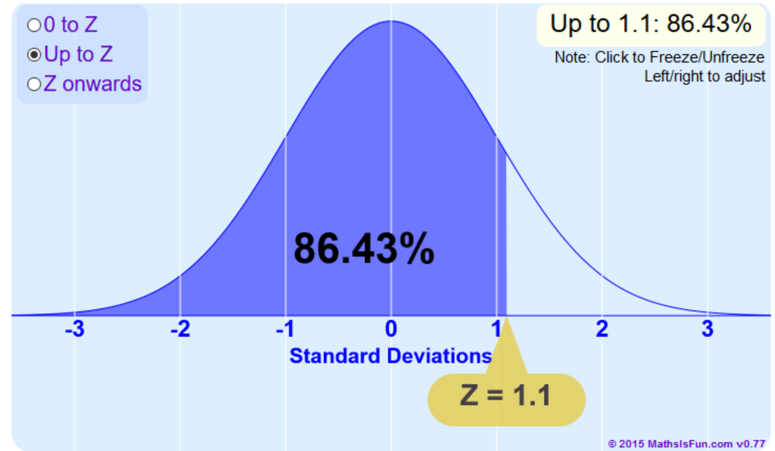
P(66" < X < 76") = P(X = 76") – P(X = 66")

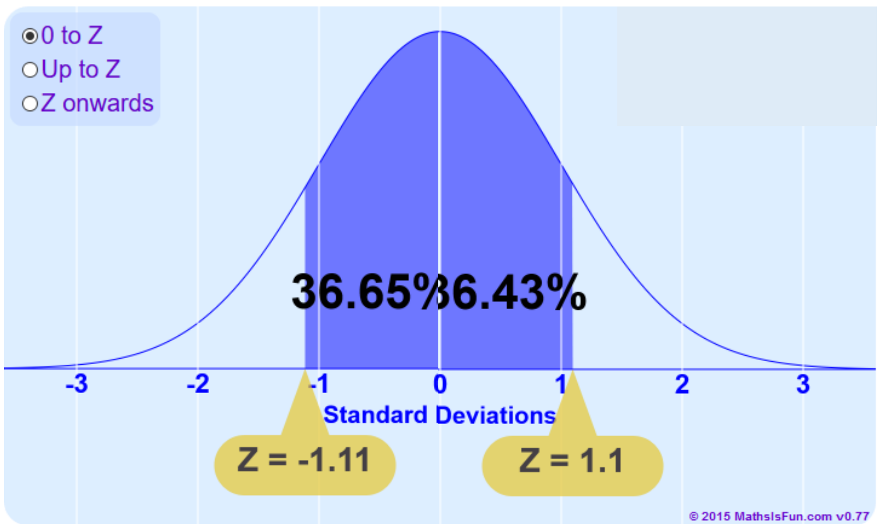
= 0.8665 – 0.1131

= 0.7534

[Z-table](https://www.mathsisfun.com/data/standard-normal-distribution-table.html)







$P(-1.11 < Z < 1.11)$

$= P(Z = 1.11) - P(Z = -1.11)$

$= 0.8665 - 0.1131$

$= 0.7534$

Attention Check

Q. What is the standard score for N(10,4), value 6?

A. $z = \frac{6-10}{2} = -2$

Q. The standard score of value 20 is 2. If the variance is 16, what is the mean?

A.

$$2 = \frac{20-\mu}{4}$$

$\therefore \mu = 20 - 8 = 12$

Attention Check

Q. Julie just realized that she wants her date to be taller when she is wearing her heels, which are 5” high. Find the new probability that her date will be taller.

A. Julie Height = 64”

$$Z = \frac{(64 + 5) - 74}{4.5} = -\ 0.44$$

$$P(Z < -0.44) = ?$$

Z-table

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
–3.4	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0002
–3.3	.0005	.0005	.0005	.0004	.0004	.0004	.0004	.0004	.0004	.0003
–3.2	.0007	.0007	.0006	.0006	.0006	.0006	.0006	.0005	.0005	.0005
–3.1	.0010	.0009	.0009	.0009	.0008	.0008	.0008	.0008	.0007	.0007
–3.0	.0013	.0013	.0013	.0012	.0012	.0011	.0011	.0011	.0010	.0010
–2.9	.0019	.0018	.0018	.0017	.0016	.0016	.0015	.0015	.0014	.0014
–2.8	.0026	.0025	.0024	.0023	.0023	.0022	.0021	.0021	.0020	.0019
–2.7	.0035	.0034	.0033	.0032	.0031	.0030	.0029	.0028	.0027	.0026
–2.6	.0047	.0045	.0044	.0043	.0041	.0040	.0039	.0038	.0037	.0036
–2.5	.0062	.0060	.0059	.0057	.0055	.0054	.0052	.0051	.0049	.0048
–2.4	.0082	.0080	.0078	.0075	.0073	.0071	.0069	.0068	.0066	.0064
–2.3	.0107	.0104	.0102	.0099	.0096	.0094	.0091	.0089	.0087	.0084
–2.2	.0139	.0136	.0132	.0129	.0125	.0122	.0119	.0116	.0113	.0110
–2.1	.0179	.0174	.0170	.0166	.0162	.0158	.0154	.0150	.0146	.0143
–2.0	.0228	.0222	.0217	.0212	.0207	.0202	.0197	.0192	.0188	.0183
–1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0239	.0233
–1.8	.0359	.0351	.0344	.0336	.0329	.0322	.0314	.0307	.0301	.0294
–1.7	.0446	.0436	.0427	.0418	.0409	.0401	.0392	.0384	.0375	.0367
–1.6	.0548	.0537	.0526	.0516	.0505	.0495	.0485	.0475	.0465	.0455
–1.5	.0668	.0655	.0643	.0630	.0618	.0606	.0594	.0582	.0571	.0559
–1.4	.0808	.0793	.0778	.0764	.0749	.0735	.0721	.0708	.0694	.0681
–1.3	.0968	.0951	.0934	.0918	.0901	.0885	.0869	.0853	.0838	.0823
–1.2	.1151	.1131	.1112	.1093	.1075	.1056	.1038	.1020	.1003	.0985
–1.1	.1357	.1335	.1314	.1292	.1271	.1251	.1230	.1210	.1190	.1170
–1.0	.1587	.1562	.1539	.1515	.1492	.1469	.1446	.1423	.1401	.1379
–0.9	.1841	.1814	.1788	.1762	.1736	.1711	.1685	.1660	.1635	.1611
–0.8	.2119	.2090	.2061	.2033	.2005	.1977	.1949	.1922	.1894	.1867
–0.7	.2420	.2389	.2358	.2327	.2296	.2266	.2236	.2206	.2177	.2148
–0.6	.2743	.2709	.2676	.2643	.2611	.2578	.2546	.2514	.2483	.2451
–0.5	.3085	.3050	.3015	.2981	.2946	.2912	.2877	.2843	.2810	.2776
–0.4	.3446	.3409	.3372	.3336	.3300	.3264	.3228	.3192	.3156	.3121
–0.3	.3821	.3783	.3745	.3707	.3669	.3632	.3594	.3557	.3520	.3483
–0.2	.4207	.4168	.4129	.4090	.4052	.4013	.3974	.3936	.3897	.3859
–0.1	.4602	.4562	.4522	.4483	.4443	.4404	.4364	.4325	.4286	.4247

INNOMATICS TECHNOLOGY HUB

Attention Check

Q. Julie just realized that she wants her date to be taller when she is wearing her heels, which are 5” high. Find the new probability that her date will be taller.

A. Julie Height = 64”

$$Z = \frac{(64+5) - 74}{4.5} = -0.44$$

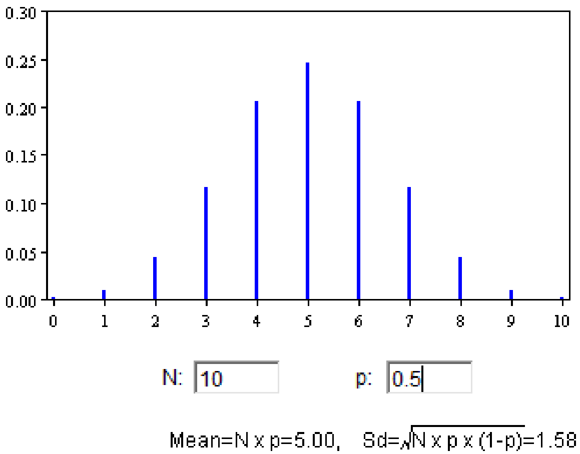
$$P(Z < -0.44) = 0.33$$

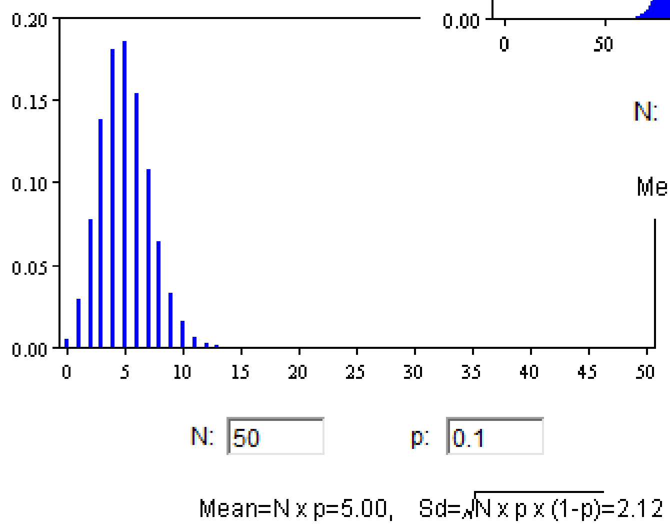
$$\therefore P(Z > -0.44) = 0.67 \text{ or } 67\%$$

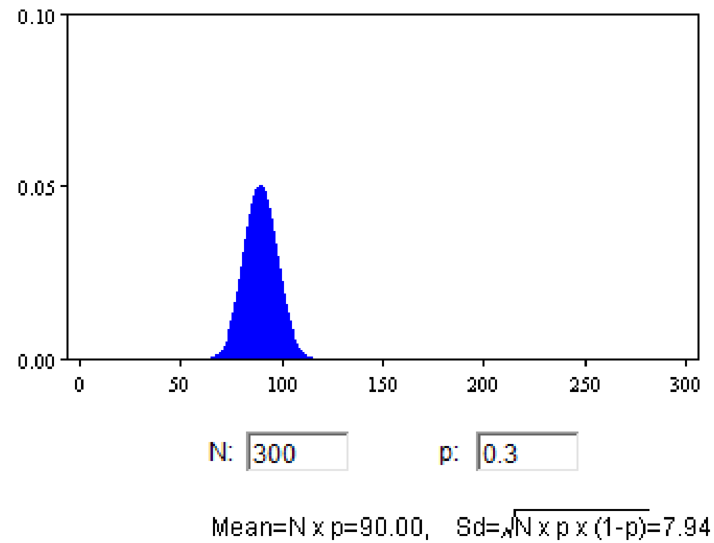
[Z-table]

Normal Distribution

Binomial distribution can be approximated to a Normal distribution if *np* > 5 and *np* > 5 (Continuity Correction required).

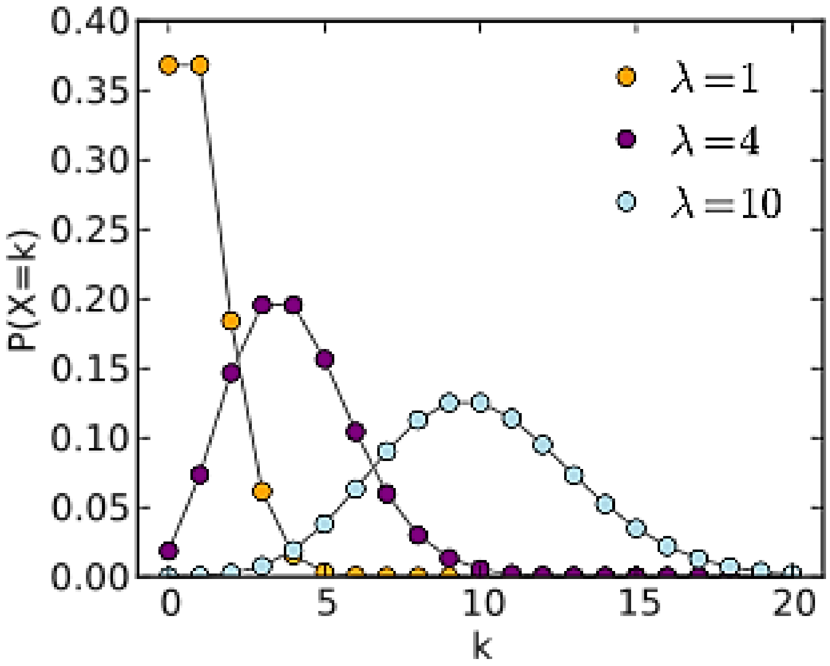






Normal distribution

Poisson distribution can be approximated to a Normal distribution when λ>15 (Continuity Correction required).



Continuity Correction?

You are playing Who Wants to Win a 

What is the probability you will get 5 or fewer correct out of 12, given that each question has only 2 possible choices? You have no lifelines.

X ~ B(12, 0.5) and we need to find P(X < 6).

$$P(X = 0) = \ {}^{12}_{0}C\ 0.5^{0}0.5^{12} = 0.5^{12}$$

$$P(X = 1) = \ {}^{12}_{1}C\ 0.5^{1}0.5^{12-1} = 12\ *\ 0.5^{12}$$

$$P(X = 2) = \ {}^{12}_{2}C\ 0.5^{2}0.5^{12-2} = 66\ *\ 0.5^{12}$$

$$P(X = 3) = \ {}^{12}_{2}C\ 0.5^{2}0.5^{12-2} = 220\ *\ 0.5^{12}$$

$$P(X = 4) = \ {}^{12}_{2}C\ 0.5^{2}0.5^{12-2} = 495\ *\ 0.5^{12}$$

$$P(X = 5) = \ {}^{12}_{2}C\ 0.5^{2}0.5^{12-2} = 792\ *\ 0.5^{12}$$

$$P(X < 6) =$$

$$(\ 1 + 12 + 66 + 220 + 495 + 792) * 0.5^{12} = 0.387$$

Continuity Correction?

X ~ B(12, 0.5) can be approximated to X ~ N(6, 3). How/Why?

n = 12, p = 0.5 and q = 0.5. Since **np** and **nq** are both > 5, the Binomial distribution can be approximated to a Normal distribution, i.e.,

X ~ B(n, p) can be approximated to X ~ N(**np**, **npq**).

If we want to get P(X < 6), what is the next step to do in the Normal distribution?

Calculate the z-score (or the standard-score).

$$Z = \frac{x - \mu}{\sigma} = \frac{6 - 6}{3} = 0$$

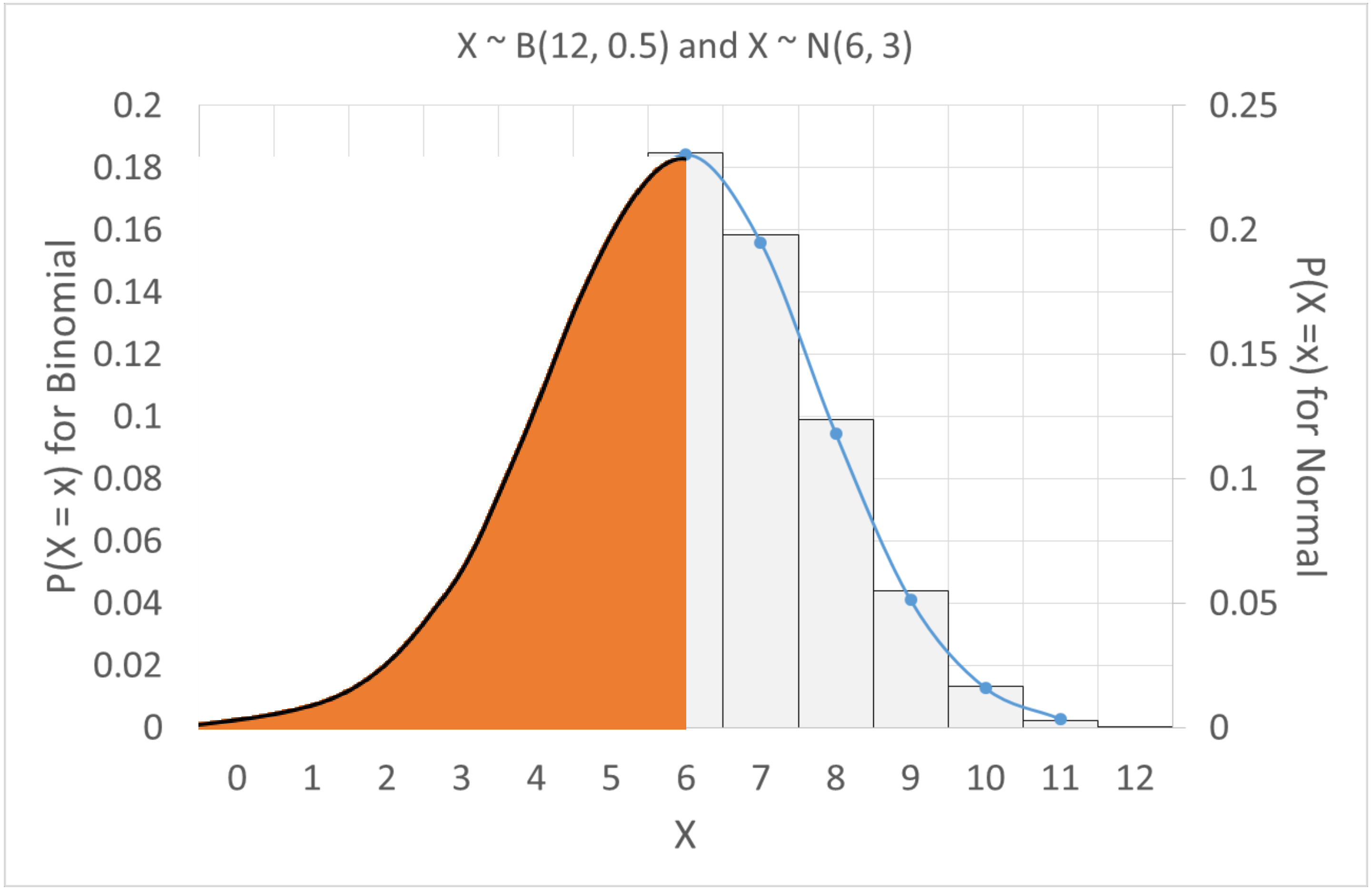
What do we do with the z-score? Look it up in the probability tables. What is the probability corresponding to the z-score of 0?

$$P(X < 6) = 0.5$$

Continuity Correction?

So, $P(X < 6) = 0.387$ for $X \sim B(12, 0.5)$

and $P(X < 6) = 0.5$ for $X \sim N(6, 3)$. Is this a good approximation?

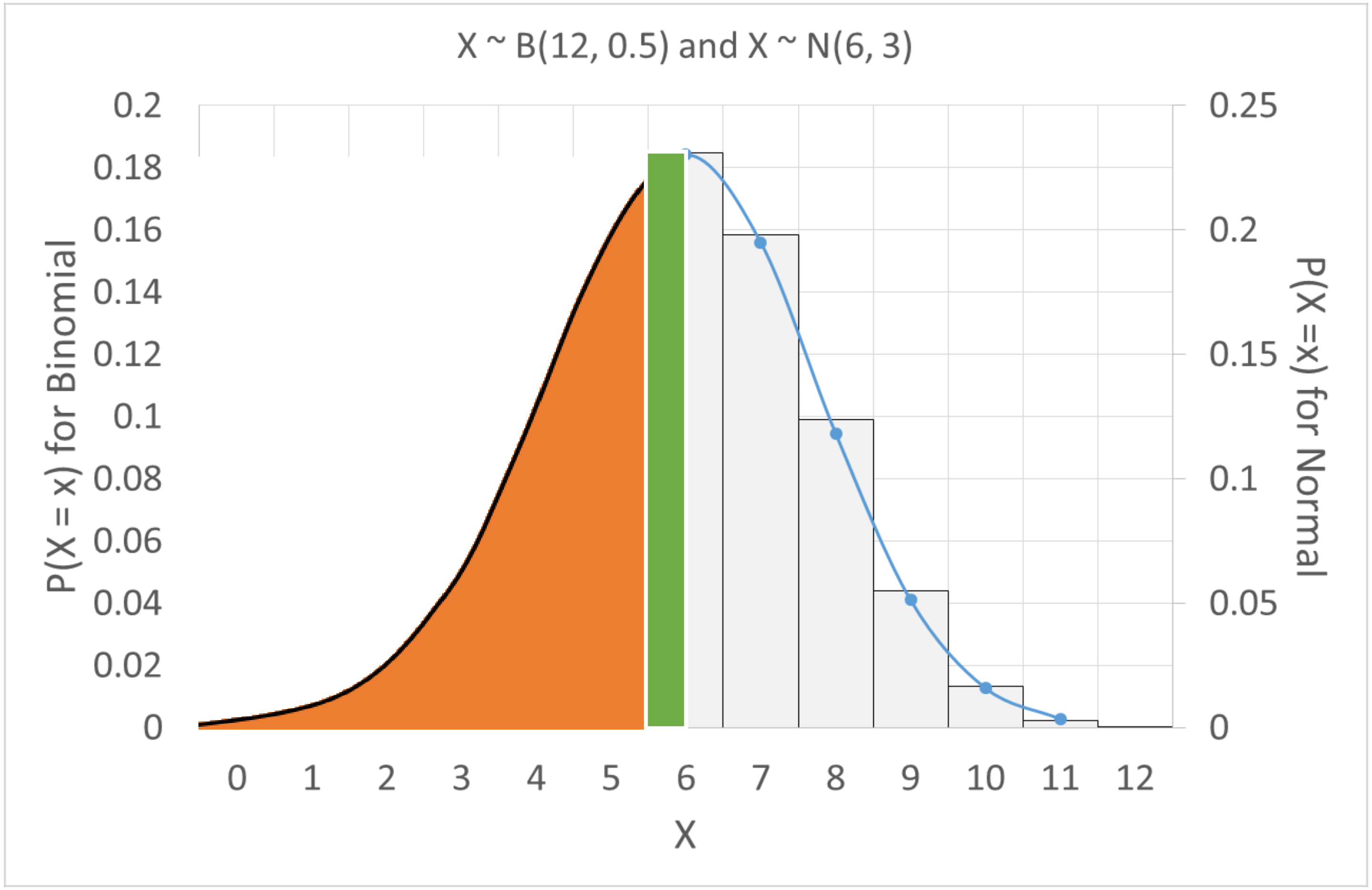


z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
0.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
0.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
0.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
0.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
0.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545

Continuity Correction?

So, $P(X < 6) = 0.387$ for $X \sim B(12, 0.5)$

and $P(X < 6) = 0.5$ for $X \sim N(6, 3)$. Is this a good approximation?



Continuity Correction?

So, P(X < 6) = 0.387 for

X ~ B(12, 0.5)

and P(X < 6) = 0.5 for

X ~ N(6, 3).

$$z = \frac{5.5 \ -6}{\sqrt{3}} = -0.29$$

P(X < 5.5) = 0.3859 for

X ~ N(6, 3)

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
–3.4	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0002
–3.3	.0005	.0005	.0005	.0004	.0004	.0004	.0004	.0004	.0004	.0003
–3.2	.0007	.0007	.0006	.0006	.0006	.0006	.0006	.0005	.0005	.0005
–3.1	.0010	.0009	.0009	.0009	.0008	.0008	.0008	.0008	.0007	.0007
–3.0	.0013	.0013	.0013	.0012	.0012	.0011	.0011	.0011	.0010	.0010
–2.9	.0019	.0018	.0018	.0017	.0016	.0016	.0015	.0015	.0014	.0014
–2.8	.0026	.0025	.0024	.0023	.0023	.0022	.0021	.0021	.0020	.0019
–2.7	.0035	.0034	.0033	.0032	.0031	.0030	.0029	.0028	.0027	.0026
–2.6	.0047	.0045	.0044	.0043	.0041	.0040	.0039	.0038	.0037	.0036
–2.5	.0062	.0060	.0059	.0057	.0055	.0054	.0052	.0051	.0049	.0048
–2.4	.0082	.0080	.0078	.0075	.0073	.0071	.0069	.0068	.0066	.0064
–2.3	.0107	.0104	.0102	.0099	.0096	.0094	.0091	.0089	.0087	.0084
–2.2	.0139	.0136	.0132	.0129	.0125	.0122	.0119	.0116	.0113	.0110
–2.1	.0179	.0174	.0170	.0166	.0162	.0158	.0154	.0150	.0146	.0143
–2.0	.0228	.0222	.0217	.0212	.0207	.0202	.0197	.0192	.0188	.0183
–1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0239	.0233
–1.8	.0359	.0351	.0344	.0336	.0329	.0322	.0314	.0307	.0301	.0294
–1.7	.0446	.0436	.0427	.0418	.0409	.0401	.0392	.0384	.0375	.0367
–1.6	.0548	.0537	.0526	.0516	.0505	.0495	.0485	.0475	.0465	.0455
–1.5	.0668	.0655	.0643	.0630	.0618	.0606	.0594	.0582	.0571	.0559
–1.4	.0808	.0793	.0778	.0764	.0749	.0735	.0721	.0708	.0694	.0681
–1.3	.0968	.0951	.0934	.0918	.0901	.0885	.0869	.0853	.0838	.0823
–1.2	.1151	.1131	.1112	.1093	.1075	.1056	.1038	.1020	.1003	.0985
–1.1	.1357	.1335	.1314	.1292	.1271	.1251	.1230	.1210	.1190	.1170
–1.0	.1587	.1562	.1539	.1515	.1492	.1469	.1446	.1423	.1401	.1379
–0.9	.1841	.1814	.1788	.1762	.1736	.1711	.1685	.1660	.1635	.1611
–0.8	.2119	.2090	.2061	.2033	.2005	.1977	.1949	.1922	.1894	.1867
–0.7	.2420	.2389	.2358	.2327	.2296	.2266	.2236	.2206	.2177	.2148
–0.6	.2743	.2709	.2676	.2643	.2611	.2578	.2546	.2514	.2483	.2451
–0.5	.3085	.3050	.3015	.2981	.2946	.2912	.2877	.2843	.2810	.2776
–0.4	.3446	.3409	.3372	.3336	.3300	.3264	.3228	.3192	.3156	.3121
–0.3	.3821	.3783	.3745	.3707	.3669	.3632	.3594	.3557	.3520	.3483
–0.2	.4207	.4168	.4129	.4090	.4052	.4013	.3974	.3936	.3897	.3859
–0.1	.4602	.4562	.4522	.4483	.4443	.4404	.4364	.4325	.4286	.4247
–0.0	.5000	.4960	.4920	.4880	.4840	.4801	.4761	.4721	.4681	.4641

Continuity Correction

Identify the right continuity correction for each discrete probability distribution.

Discrete	Continuous
$X < 3$	$X < 2.5$
$X > 3$	$X > 3.5$
$X \leq 3$	$X < 3.5$
$X \geq 3$	$X > 2.5$
$3 \leq X < 10$	$2.5 < X < 9.5$
$X = 0$	$-0.5 < X < 0.5$
$3 \leq X \leq 10$	$2.5 < X < 10.5$
$X > 0$	$X > 0.5$
$3 < X < 10$	$2.5 < X < 9.5$

Continuity Correction?

You are playing Who Wants to Win a 

What is the probability of getting 30 out of 40 questions correct, where each question has possible choices ?

X ~ B(40, 0.5) and we need to find P(X $\geq$ 30).

P(X $>$ 29.5) where X ~ N(20,10).

$$Z = \frac{29.5 \ - 20}{\sqrt{10}} = 3.004$$

$$P(\,X > 29.5) = 1 \ - 0.9987 = \ \ 0.0013$$

Z-table

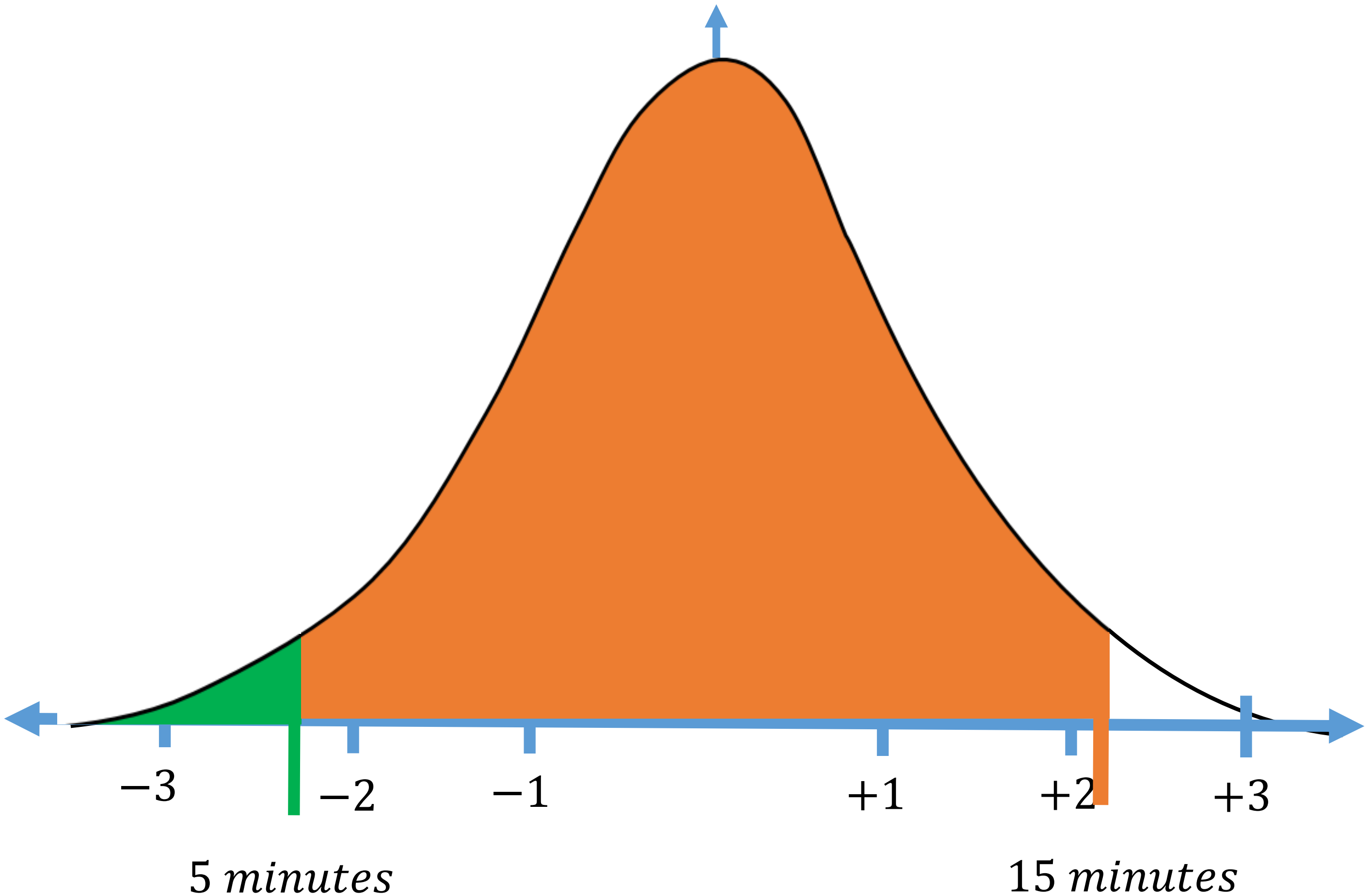
Normal Distribution

You have designed a new game, Angry Birds. The key to success is that it should not be so difficult that people get frustrated, nor should it be so easy that they don't get challenged. Before building the new level, you want to know what the mean and standard deviation are of the number of minutes people take to complete level 1.

You know the following:

1. The # of minutes follows a normal distribution.
2. The probability of a player playing for less than 5 minutes is 0.0045.
3. The probability of a player playing for less than 15 minutes is 0.9641.

Normal Distribution



Normal Distribution

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
−3.4	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0002
−3.3	.0005	.0005	.0005	.0004	.0004	.0004	.0004	.0004	.0004	.0003
−3.2	.0007	.0007	.0006	.0006	.0006	.0006	.0006	.0005	.0005	.0005
−3.1	.0010	.0009	.0009	.0009	.0008	.0008	.0008	.0008	.0007	.0007
−3.0	.0013	.0013	.0013	.0012	.0012	.0011	.0011	.0011	.0010	.0010
−2.9	.0019	.0018	.0018	.0017	.0016	.0016	.0015	.0015	.0014	.0014
−2.8	.0026	.0025	.0024	.0023	.0023	.0022	.0021	.0021	.0020	.0019
−2.7	.0035	.0034	.0033	.0032	.0031	.0030	.0029	.0028	.0027	.0026
−2.6	.0047	.0045	.0044	.0043	.0041	.0040	.0039	.0038	.0037	.0036
−2.5	.0062	.0060	.0059	.0057	.0055	.0054	.0052	.0051	.0049	.0048
−2.4	.0082	.0080	.0078	.0075	.0073	.0071	.0069	.0068	.0066	.0064
−2.3	.0107	.0104	.0102	.0099	.0096	.0094	.0091	.0089	.0087	.0084
−2.2	.0139	.0136	.0132	.0129	.0125	.0122	.0119	.0116	.0113	.0110
−2.1	.0179	.0174	.0170	.0166	.0162	.0158	.0154	.0150	.0146	.0143
−2.0	.0228	.0222	.0217	.0212	.0207	.0202	.0197	.0192	.0188	.0183
−1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0239	.0233
−1.8	.0359	.0351	.0344	.0336	.0329	.0322	.0314	.0307	.0301	.0294
−1.7	.0446	.0436	.0427	.0418	.0409	.0401	.0392	.0384	.0375	.0367
−1.6	.0548	.0537	.0526	.0516	.0505	.0495	.0485	.0475	.0465	.0455
−1.5	.0668	.0655	.0643	.0630	.0618	.0606	.0594	.0582	.0571	.0559
−1.4	.0808	.0793	.0778	.0764	.0749	.0735	.0721	.0708	.0694	.0681
−1.3	.0968	.0951	.0934	.0918	.0901	.0885	.0869	.0853	.0838	.0823
−1.2	.1151	.1131	.1112	.1093	.1075	.1056	.1038	.1020	.1003	.0985
−1.1	.1357	.1335	.1314	.1292	.1271	.1251	.1230	.1210	.1190	.1170
−1.0	.1587	.1562	.1539	.1515	.1492	.1469	.1446	.1423	.1401	.1379
−0.9	.1841	.1814	.1788	.1762	.1736	.1711	.1685	.1660	.1635	.1611
−0.8	.2119	.2090	.2061	.2033	.2005	.1977	.1949	.1922	.1894	.1867
−0.7	.2420	.2389	.2358	.2327	.2296	.2266	.2236	.2206	.2177	.2148
−0.6	.2743	.2709	.2676	.2643	.2611	.2578	.2546	.2514	.2483	.2451

$P(X < 5) = 0.0045$

$Z_1 = -2.61$

Normal Distribution

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
0.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
0.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
0.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
0.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
0.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817

$P(X < 15)$
$= 0.9641$

$Z_2 = 1.8$

Normal Distribution

$-2.61 = \frac{5-\mu}{\sigma}$ and $1.8 = \frac{15-\mu}{\sigma}$

Solving for the above 2 equations, we get

$\mu = 5 + 2.61\sigma$

$\mu = 15 - 1.8\sigma$

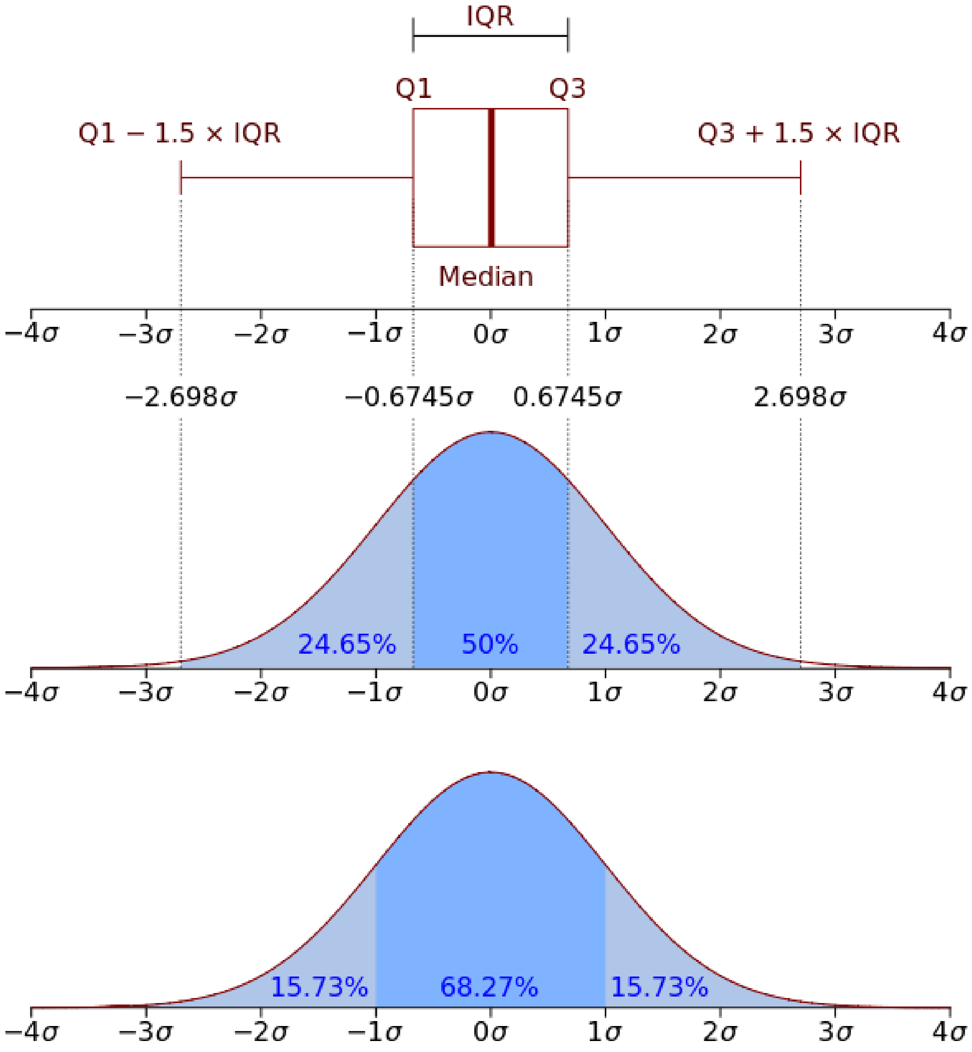
Subtracting the two, we get

$0 = -10 + 4.41\sigma \Rightarrow \sigma = 10 \div 4.41 = 2.27$

Substituting this value of σ in either of the above 2 equations,

we get $\mu = 5 + 2.61 * 2.27 = 10.925$

Z-table



IQR
Q1
Q3
Q1 − 1.5 × IQR
Q3 + 1.5 × IQR
Median
−4σ
−3σ
−2σ
−1σ
0σ
1σ
2σ
3σ
4σ
−2.698σ
−0.6745σ
0.6745σ
2.698σ
24.65%
50%
24.65%
15.73%
68.27%
15.73%

INNOMATICS TECHNOLOGY HUB

A sister concern of

KenMinds
...committed to solution excellence