Editing file

Feedback

Normal distribution and its application

Objectives:

1. Able to understand the concept of Normal distribution.
2. Able to calculate the z-score for quantitative variable.
3. Able to apply the concept in the interpretation of a clinical data.

Click here for the practical

24th lecture

Color Index:

- Boys' Slides
- Girls' Slides
- Doctors Notes
- Golden Notes
- Important
- Extra

Revised by

MED439 King Saud University

Research Team

RESEARCH TEAM 438

overview

Gaussian Distribution[1]:

The Normal or Gaussian distribution is the most important continuous probability distribution in statistics.

1- The term “Gaussian” refers to ‘Carl Freidrich Gauss’ who develop this distribution.

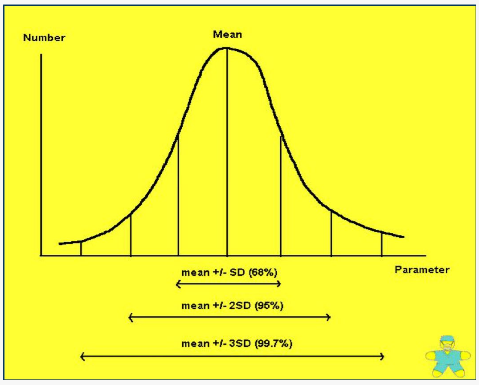
2- The word ‘normal’ here means that the distribution confirms to a certain formula and shape.(does not mean ‘ordinary’ or ‘common’ nor does it mean ‘disease-free’)

3- Many biologic variables follow this pattern eg.(Hemoglobin, Cholesterol, Serum Electrolytes, Blood pressures, age, weight, height)[2].

4- One can use this information to define what is normal and what is extreme.

5- In clinical medicine 95% or 2 Standard deviations around the mean is normal.[3]

6- Clinically, 5% of “normal” individuals are labeled as extreme/abnormal.
→ We just accept this and move on.



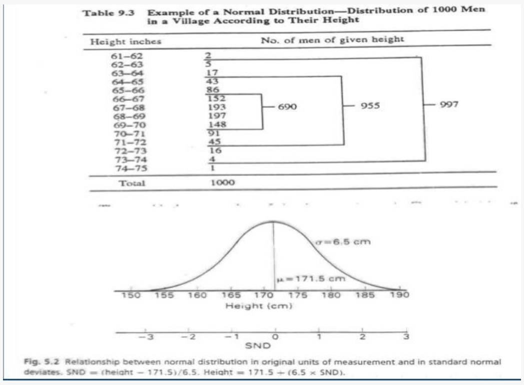


Table 9.3 Example of a Normal Distribution—Distribution of 1000 Men in a Village According to Their Height

Height inches	No. of men of given height
61–62	2
62–63	5
63–64	17
64–65	43
65–66	86
66–67	152
67–68	193
68–69	197
69–70	148
70–71	91
71–72	45
72–73	16
73–74	4
74–75	1
Total	1000

Fig. 5.2 Relationship between normal distribution in original units of measurement and in standard normal deviates. SND = (height − 171.5)/6.5. Height = 171.5 + (6.5 × SND).

Characteristics of Normal Distribution:

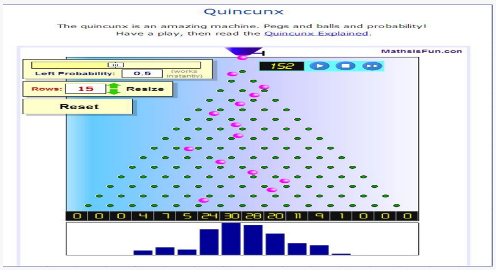
- Symmetrical about mean, (μ).
- Mean, median, and mode are equal.
- Total area under the curve above the x-axis is one square unit. Which is 100% or probability 1.
- 1 standard deviation on both sides of the mean includes approximately 68% of the total area
- 2 standard deviations on both sides of the mean includes approximately 95% of the total area
- 3 standard deviations on both sides of the mean includes approximately 99% of the total area

1) In statistics, there are more than 1300 type of distribution. You are studying one type which is normal distribution.
2) All are Quantitative continuous/Quantitative discrete values
3) How do they set normal range in medicine? They accept two SD around the mean as “normal”.

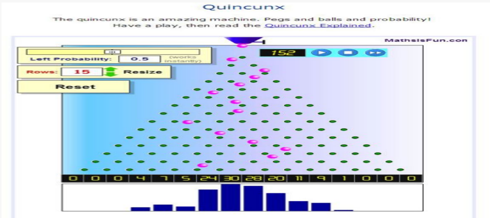
Cont..

Pascal Triangle (Quincunx)

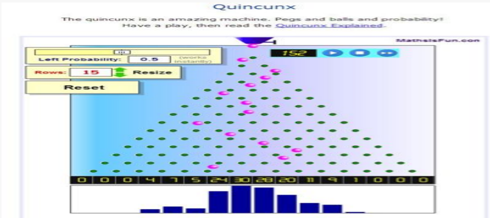
- In the past this experiment was done in a physical box, but today we can use a computer simulation.
- Nails were punched into a box to form a triangular shape.



- On top there is only one nail. The second row has two nails. Each subsequent row has one additional nail
- When a ball is poured into the box from top and lands on the first nail, the probability of going to the left is .5 and to the right is also .5.

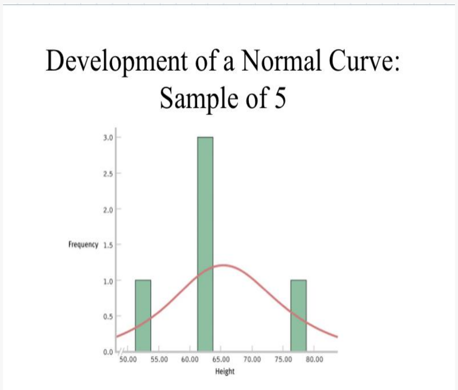


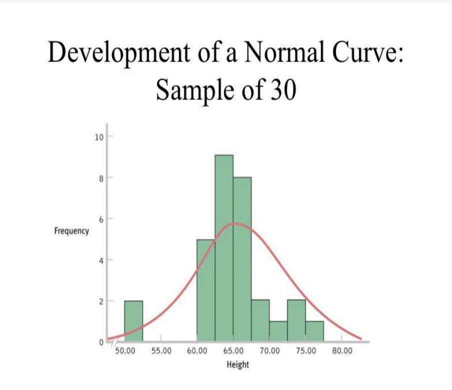
- Subsequently, the probability of going to which direction gets more and more complicated. Nonetheless, the process is random.
- But this random process always produces a normal distribution!
- http://www.mathsisfun.com/data/quincunx.html

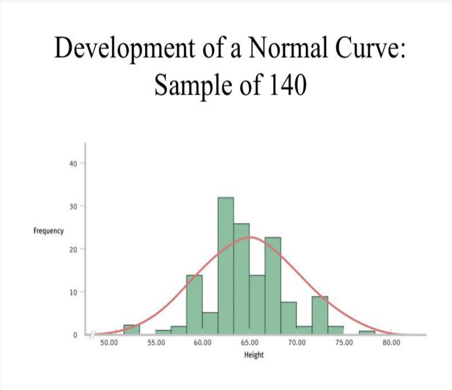


Development of a normal curve

Sample size must be large (minimal is 30.00) any data that less than 30.00 we won't consider it as normal distribution







When the sample size increase > the shape getting better

Cont..

Uses of Normal Distribution:

1- It's application goes beyond describing distributions.
2- It is used by researchers.
3- The major use of normal distribution is the role it plays in statistical inference
4- It helps managers to make decisions.

What's so Great about the Normal Distribution?

- If you know two things,
  - → Mean.
  - → Standard deviation.
- Then you know everything about the distribution, and you know the probability of any value arising.

Standardised score (z-score):

My diastolic blood pressure is 100. So what?

Normal is 90 (for my age and sex). Mine is high, but how much high?

Express it in standardised scores. How many SDs above the mean is that?

Mean = 90, SD = 4 (my age and sex).

- My score - Mean score /SD = 100-90 /4 =2.5
- This is a standardised score, or z-score.
- Look z tables (or computer) .
  -See how often this high (or higher) score occur.

As we discussed earlier, Two SD around mean is considered normal , but this blood pressure is 2.5 above mean so it is abnormal and has to be treated.

Z-Score:

- When a set of data values are normally distributed, we can standardize each score by converting it into a z-score.
- z-scores make it easier to compare data values measured on different scales. example: (age and years/ BMI and other units)
- A z-score reflects how many standard deviations above or below the mean a raw score is.
- The z-score is positive if the data value lies above the mean and negative if the data value lies below the mean.

Standardised score (z-score)

Z-Score:

⊙ The Z-score makes it possible, under some circumstances, to compare scores that originally had different units of measurement

Comparing Apples and Oranges

If we can standardize the raw scores on two different scales, converting both scores to z scores, we can then compare the scores directly.

Using z Scores to Make Comparisons

- If you know your score on an exam, and a friend's score on an exam, you can convert to z scores to determine who did better and by how much.
- z scores are standardized, so they can be compared!

Suppose you scored a 60 on a numerical test and a 30 on a verbal test:

⊙ On which test did you perform better?

First, we need to know how other people did on the same tests.

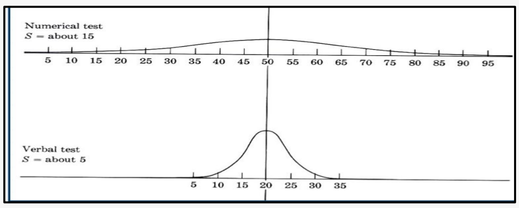
→ Suppose that the mean score on the numerical test was 50 and the mean score on the verbal test was 20. You scored 10 points above the mean on each test.

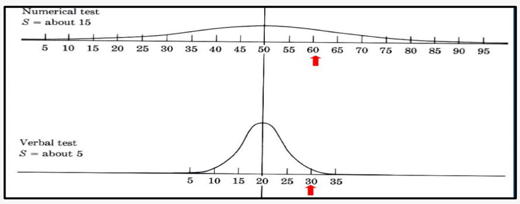
⊙ Can you conclude that you did equally well on both tests?

You do not know, because you do not know if 10 points on the numerical test is the same as 10 points on the verbal test.

→ Suppose also that the standard deviation on the numerical test was 15 and the standard deviation on the verbal test was 5.

⊙ Now can you determine on which test you did better? ○ Yes, check the page below to know HOW!





⊙The larger the SD the more flat the normal distribution becomes.

⊙ To find out how many standard deviations away from the mean a particular score is, use the Z formula: standard deviation is the unit of measurement

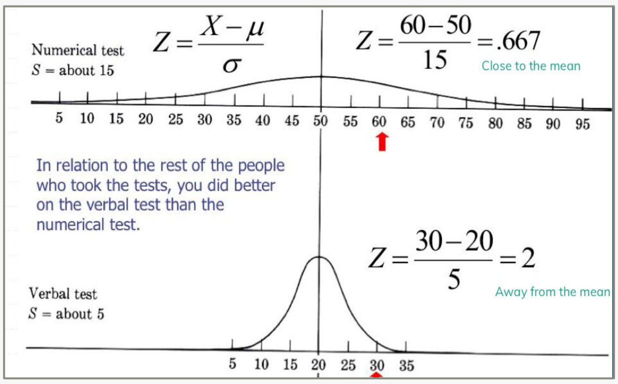
Where x represents an element of the data set, the mean is represented by and standard deviation by .

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

Population

$$Z = \frac{X - \bar{X}}{S}$$

Sample



Performed better in the verbal test because he scored 2 SD above the average in comparison to only 0.6 SD in the numerical test

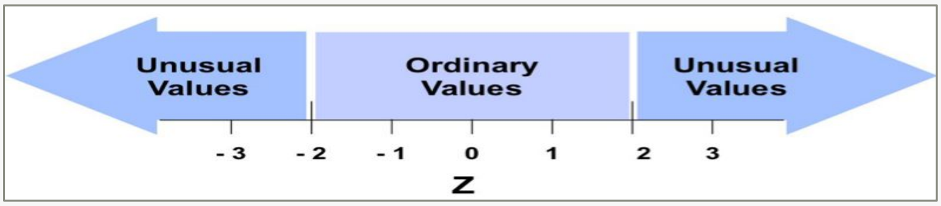
Interpretation

Properties of Z-score:

⊙ Allows you to describe a particular score in terms of where it fits into the overall group of scores.
-Whether it is above or below the average and how much it is above or below the average.

⊙ A standard score that states the position of a score in relation to the mean of the distribution, using the standard deviation as the unit of measurement.
-The number of standard deviations a score is above or below a mean.

Interpreting Z Scores:



The Standard Normal Table:

⊙ Using the standard normal table, you can find the area under the curve that corresponds with certain scores.

⊙ The area under the curve is proportional to the frequency of scores.

⊙ The area under the curve gives the probability of that score occurring.

The Tables:

When you have a Z score, and you want to get the probability of that Z score, you go to the table and find where exactly your Z value fall in. if your Z value is negative (below the mean) you can use the same table here as the normal distribution is symmetrical on both sides so no need for another table.

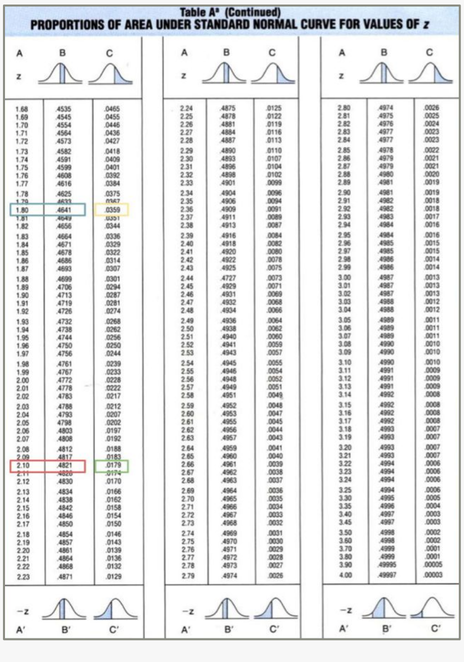
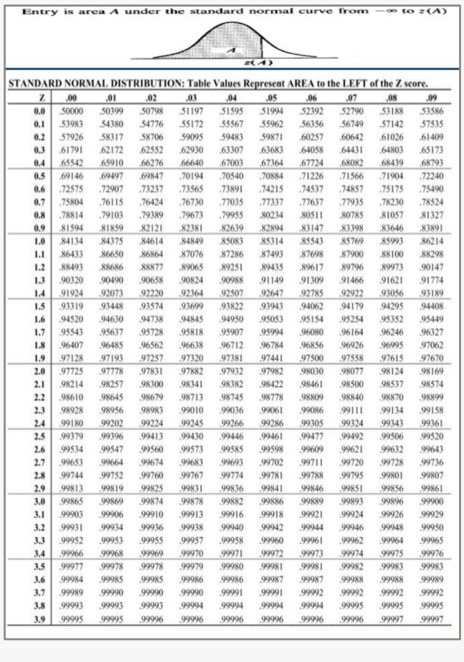


Table A* (Continued)
PROPORTIONS OF AREA UNDER STANDARD NORMAL CURVE FOR VALUES OF z

A	B	C	A	B	C	A	B	C
z			z			z		
1.68	.4535	.0465	2.24	.4875	.0125	2.80	.4974	.0026
1.69	.4545	.0455	2.25	.4878	.0122	2.81	.4975	.0025
1.70	.4554	.0446	2.26	.4881	.0119	2.82	.4976	.0024
1.71	.4564	.0436	2.27	.4884	.0116	2.83	.4977	.0023
1.72	.4573	.0427	2.28	.4887	.0113	2.84	.4977	.0023
1.73	.4582	.0418	2.29	.4890	.0110	2.85	.4978	.0022
1.74	.4591	.0409	2.30	.4893	.0107	2.86	.4979	.0021
1.75	.4599	.0401	2.31	.4896	.0104	2.87	.4979	.0021
1.76	.4608	.0392	2.32	.4898	.0102	2.88	.4980	.0020
1.77	.4616	.0384	2.33	.4901	.0099	2.89	.4981	.0019
1.78	.4625	.0375	2.34	.4904	.0096	2.90	.4981	.0019
1.79	.4633	.0367	2.35	.4906	.0094	2.91	.4982	.0018
1.80	.4641	.0359	2.36	.4909	.0091	2.92	.4982	.0018
1.81	.4649	.0351	2.37	.4911	.0089	2.93	.4983	.0017
1.82	.4656	.0344	2.38	.4913	.0087	2.94	.4984	.0016
1.83	.4664	.0336	2.39	.4916	.0084	2.95	.4984	.0016
1.84	.4671	.0329	2.40	.4918	.0082	2.96	.4985	.0015
1.85	.4678	.0322	2.41	.4920	.0080	2.97	.4985	.0015
1.86	.4686	.0314	2.42	.4922	.0078	2.98	.4986	.0014
1.87	.4693	.0307	2.43	.4925	.0075	2.99	.4986	.0014
1.88	.4699	.0301	2.44	.4727	.0073	3.00	.4987	.0013
1.89	.4706	.0294	2.45	.4929	.0071	3.01	.4987	.0013
1.90	.4713	.0287	2.46	.4931	.0069	3.02	.4987	.0013
1.91	.4719	.0281	2.47	.4932	.0068	3.03	.4988	.0012
1.92	.4726	.0274	2.48	.4934	.0066	3.04	.4988	.0012
1.93	.4732	.0268	2.49	.4936	.0064	3.05	.4989	.0011
1.94	.4738	.0262	2.50	.4938	.0062	3.06	.4989	.0011
1.95	.4744	.0256	2.51	.4940	.0060	3.07	.4989	.0011
1.96	.4750	.0250	2.52	.4941	.0059	3.08	.4990	.0010
1.97	.4756	.0244	2.53	.4943	.0057	3.09	.4990	.0010
1.98	.4761	.0239	2.54	.4945	.0055	3.10	.4990	.0010
1.99	.4767	.0233	2.55	.4946	.0054	3.11	.4991	.0009
2.00	.4772	.0228	2.56	.4948	.0052	3.12	.4991	.0009
2.01	.4778	.0222	2.57	.4949	.0051	3.13	.4991	.0009
2.02	.4783	.0217	2.58	.4951	.0049	3.14	.4992	.0008
2.03	.4788	.0212	2.59	.4952	.0048	3.15	.4992	.0008
2.04	.4793	.0207	2.60	.4953	.0047	3.16	.4992	.0008
2.05	.4798	.0202	2.61	.4955	.0045	3.17	.4992	.0008
2.06	.4803	.0197	2.62	.4956	.0044	3.18	.4993	.0007
2.07	.4808	.0192	2.63	.4957	.0043	3.19	.4993	.0007
2.08	.4812	.0188	2.64	.4959	.0041	3.20	.4993	.0007
2.09	.4817	.0183	2.65	.4960	.0040	3.21	.4993	.0007
2.10	.4821	.0179	2.66	.4961	.0039	3.22	.4994	.0006
2.11	.4826	.0174	2.67	.4962	.0038	3.23	.4994	.0006
2.12	.4830	.0170	2.68	.4963	.0037	3.24	.4994	.0006
2.13	.4834	.0166	2.69	.4964	.0036	3.25	.4994	.0006
2.14	.4838	.0162	2.70	.4965	.0035	3.30	.4995	.0005
2.15	.4842	.0158	2.71	.4966	.0034	3.35	.4996	.0004
2.16	.4846	.0154	2.72	.4967	.0033	3.40	.4997	.0003
2.17	.4850	.0150	2.73	.4968	.0032	3.45	.4997	.0003
2.18	.4854	.0146	2.74	.4969	.0031	3.50	.4998	.0002
2.19	.4857	.0143	2.75	.4970	.0030	3.60	.4998	.0002
2.20	.4861	.0139	2.76	.4971	.0029	3.70	.4999	.0001
2.21	.4864	.0136	2.77	.4972	.0028	3.80	.4999	.0001
2.22	.4868	.0132	2.78	.4973	.0027	3.90	.49995	.00005
2.23	.4871	.0129	2.79	.4974	.0026	4.00	.49997	.00003
−z			−z			−z		
A′	B′	C′	A′	B′	C′	A′	B′	C′



Entry is area A under the standard normal curve from $-\infty$ to $z(A)$

STANDARD NORMAL DISTRIBUTION: Table Values Represent AREA to the LEFT of the Z score.

Z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.50000	.50399	.50798	.51197	.51595	.51994	.52392	.52790	.53188	.53586
0.1	.53983	.54380	.54776	.55172	.55567	.55962	.56356	.56749	.57142	.57535
0.2	.57926	.58317	.58706	.59095	.59483	.59871	.60257	.60642	.61026	.61409
0.3	.61791	.62172	.62552	.62930	.63307	.63683	.64058	.64431	.64803	.65173
0.4	.65542	.65910	.66276	.66640	.67003	.67364	.67724	.68082	.68439	.68793
0.5	.69146	.69497	.69847	.70194	.70540	.70884	.71226	.71566	.71904	.72240
0.6	.72575	.72907	.73237	.73565	.73891	.74215	.74537	.74857	.75175	.75490
0.7	.75804	.76115	.76424	.76730	.77035	.77337	.77637	.77935	.78230	.78524
0.8	.78814	.79103	.79389	.79673	.79955	.80234	.80511	.80785	.81057	.81327
0.9	.81594	.81859	.82121	.82381	.82639	.82894	.83147	.83398	.83646	.83891
1.0	.84134	.84375	.84614	.84849	.85083	.85314	.85543	.85769	.85993	.86214
1.1	.86433	.86650	.86864	.87076	.87286	.87493	.87698	.87900	.88100	.88298
1.2	.88493	.88686	.88877	.89065	.89251	.89435	.89617	.89796	.89973	.90147
1.3	.90320	.90490	.90658	.90824	.90988	.91149	.91309	.91466	.91621	.91774
1.4	.91924	.92073	.92220	.92364	.92507	.92647	.92785	.92922	.93056	.93189
1.5	.93319	.93448	.93574	.93699	.93822	.93943	.94062	.94179	.94295	.94408
1.6	.94520	.94630	.94738	.94845	.94950	.95053	.95154	.95254	.95352	.95449
1.7	.95543	.95637	.95728	.95818	.95907	.95994	.96080	.96164	.96246	.96327
1.8	.96407	.96485	.96562	.96638	.96712	.96784	.96856	.96926	.96995	.97062
1.9	.97128	.97193	.97257	.97320	.97381	.97441	.97500	.97558	.97615	.97670
2.0	.97725	.97778	.97831	.97882	.97932	.97982	.98030	.98077	.98124	.98169
2.1	.98214	.98257	.98300	.98341	.98382	.98422	.98461	.98500	.98537	.98574
2.2	.98610	.98645	.98679	.98713	.98745	.98778	.98809	.98840	.98870	.98899
2.3	.98928	.98956	.98983	.99010	.99036	.99061	.99086	.99111	.99134	.99158
2.4	.99180	.99202	.99224	.99245	.99266	.99286	.99305	.99324	.99343	.99361
2.5	.99379	.99396	.99413	.99430	.99446	.99461	.99477	.99492	.99506	.99520
2.6	.99534	.99547	.99560	.99573	.99585	.99598	.99609	.99621	.99632	.99643
2.7	.99653	.99664	.99674	.99683	.99693	.99702	.99711	.99720	.99728	.99736
2.8	.99744	.99752	.99760	.99767	.99774	.99781	.99788	.99795	.99801	.99807
2.9	.99813	.99819	.99825	.99831	.99836	.99841	.99846	.99851	.99856	.99861
3.0	.99865	.99869	.99874	.99878	.99882	.99886	.99889	.99893	.99896	.99900
3.1	.99903	.99906	.99910	.99913	.99916	.99918	.99921	.99924	.99926	.99929
3.2	.99931	.99934	.99936	.99938	.99940	.99942	.99944	.99946	.99948	.99950
3.3	.99952	.99953	.99955	.99957	.99958	.99960	.99961	.99962	.99964	.99965
3.4	.99966	.99968	.99969	.99970	.99971	.99972	.99973	.99974	.99975	.99976
3.5	.99977	.99978	.99978	.99979	.99980	.99981	.99981	.99982	.99983	.99983
3.6	.99984	.99985	.99985	.99986	.99986	.99987	.99987	.99988	.99988	.99989
3.7	.99989	.99990	.99990	.99990	.99991	.99991	.99992	.99992	.99992	.99992
3.8	.99993	.99993	.99993	.99994	.99994	.99994	.99994	.99995	.99995	.99995
3.9	.99995	.99995	.99996	.99996	.99996	.99996	.99996	.99996	.99997	.99997

Interpretation

Reading Z-Tables:

Pay attention to the color code to find the answers in the table above

Finding the proportion of observations between the mean and a score when
Z=1.80

Finding the proportion of observations above a score when
Z=1.80

Finding the proportion of observations between a score and the mean when
Z=-2.10

Finding the proportion of observations below a score when
Z=-2.10

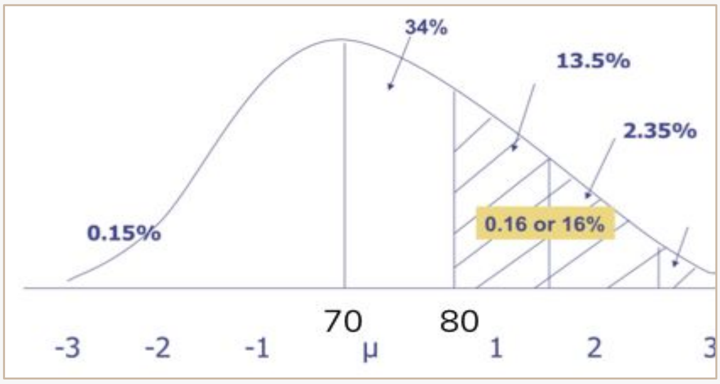
Z scores and the Normal Distribution:

⊙ They can answer a wide variety of questions about any normal distribution with a known mean and standard deviation.

⊙ Will address how to solve two main types of normal curve problems:
-Finding a proportion given a score.
-Finding a score given a proportion. This example i'm not having here, you don't need to worry about it.

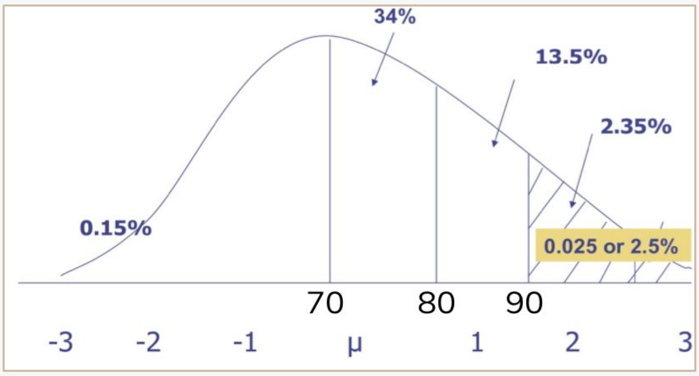
Exercise

Problem: Assuming the normal heart rate (H.R) in normal healthy individuals is normally distributed with Mean = 70 and Standard Deviation =10 beats/min.
Then:
1. What area under the curve is above 80 beats/min? Ans: 0.16 (16%).

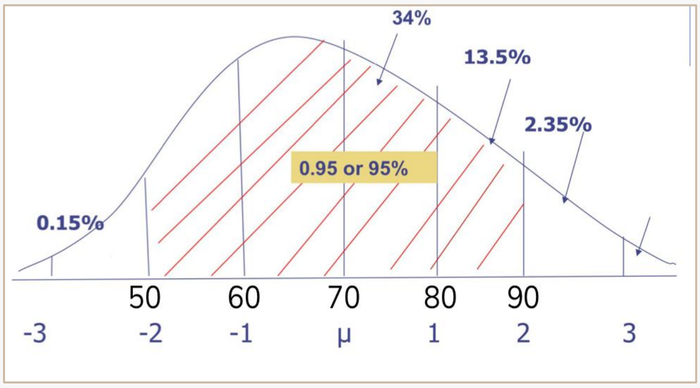


Exercises

2. What area of the curve is above 90 beats/min? Ans: 0.025 (2.5%).

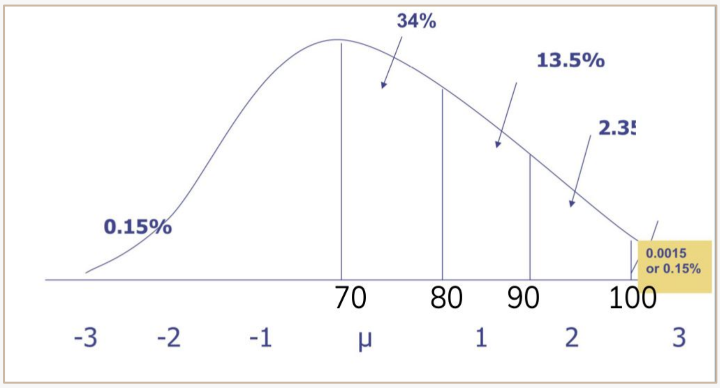


3. What area of the curve is between 50-90 beats/min? Ans: 0.95 (95%).

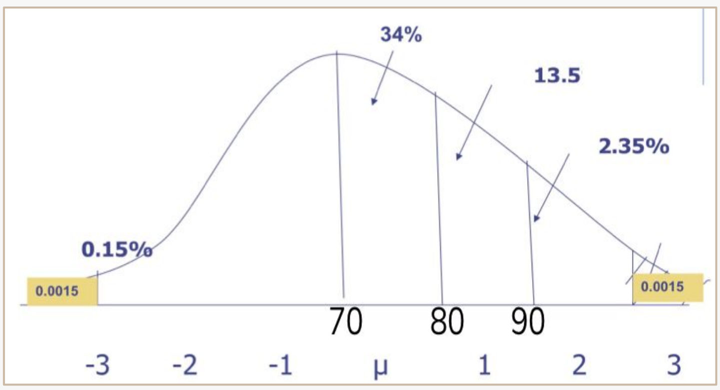


2 SD around the mean (on both sides) so it's 95% as we discussed

4. What area of the curve is above 100 beats/min? Ans: 0.0015 (0.15%).



5. What area of the curve is below 40 beats per min or above 100 beats per min? Ans: 0.0015 for each tail or 0.3%.



Exercises

Problem: Assume that among diabetics the fasting blood level of glucose is approximately normally distributed with a mean of 105 mg per 100ml and an SD of 9 mg per 100 ml. What proportion of diabetics having fasting blood glucose levels between 90 and 125 mg per 100 ml?

- Let X be the random variable denoting the fasting blood glucose level.
- X has a normal distribution with mean= 105, and SD= 9.
- We have to compute P(90≤ X≤ 125).
- The table is available only for the probabilities of a standard normal distribution. Thus we have to convert X to a standard normal variable (Z).
- Using the formula on the fourth slide of this lecture, P(90≤ X≤ 125) can be written as:
  - $P\left[\frac{90-105}{9} \leq \frac{X-105}{9} \leq \frac{125-105}{9}\right] = P(-1.67 \leq Z \leq 2.22)$
  - Since $Z = \frac{X-105}{9}$
  - P(Z≤2.22) - P(Z<-1.67) = 0.9868-0.0475 = 0.9393.
  - Therefore 94% of diabetics have fasting blood glucose levels between 90 and 125.

Another example: The mean marks of the class is 75 and SD is 5. What proportion of students' marks are there between between 60-80? What proportion of students scored greater than 65?
So you'll get three scenarios:
1) In between
2) Less than
3) Greater than
The normal distribution is found to answer these type of questions.

Dr notes

- In the exam i'm not going to give Z score fractions and no need to worry about the table, if I give this I have to provide the normal distribution table in the exam and that would be a problem. Main thing is the concept of Z score for example:
  The mean Hb of a particular group is X and the SD is Y, what is the Z value? You can use the simple expression and get the Z value.
- Concentrate on the characteristics of normal distribution

Lecture Summary

Normal Distribution:

The Normal or Gaussian distribution is the most important continuous probability distribution in statistics

Many biologic variables follow this pattern

The role it plays in statistical inference helps managers to make decisions.

Total area under the curve above the x-axis is one square unit.

Total area under the curve above the x-axis is one square unit.

Symmetrical about mean, (μ).

Standardised score (z-score)and the standard normal table:

Z-Score (or standard score): is the number of standard deviations that a given value x is above or below the mean.

The Z-score makes it possible, under some circumstances, to compare scores that originally had different units of measurement.

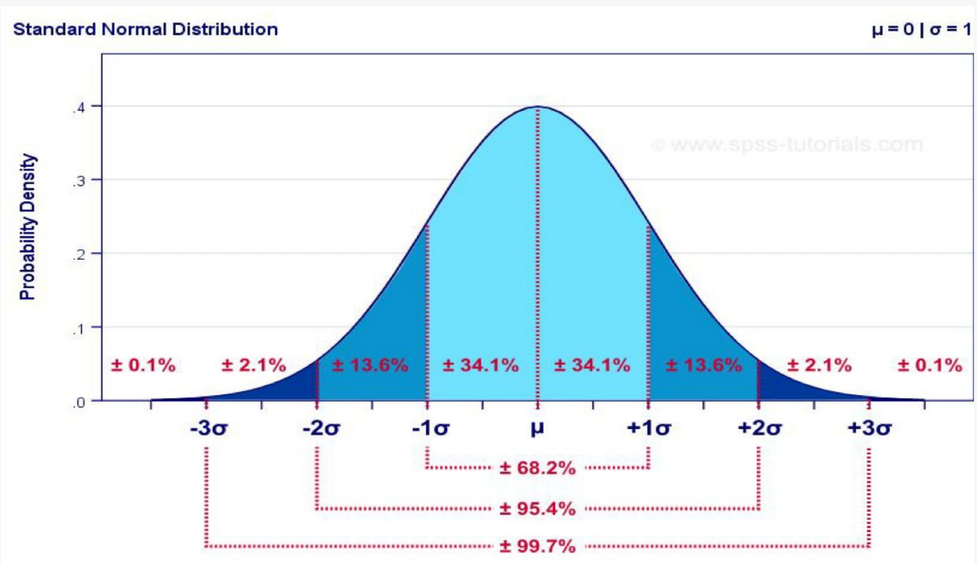
$$Z = \frac{X - \mu}{\sigma}$$

Using the standard normal table, you can find the area under the curve that corresponds with certain scores.

The area under the curve is proportional to the frequency of scores.

The area under the curve is gives the probability of that score occurring.

This is the standard normal curve with all percentages.



Questions

(1) In Standard normal distribution, the value of mode is?

A) 1

B) 0

C) 2

D) Not fixed

(2) The area under a standard normal curve is?

A) 0

B) 1

C) ∞

D) not defined

(3) For a standard normal variate, the value of mean is?

A) 0

B) 1

C) ∞

D) not defined.

(4) Normal Distribution is applied for:

A) Continuous Random Distribution.

B) Discrete Random Variable.

C) Irregular Random Variable.

D) Uncertain Random Variable.

(5) Normal Distribution is symmetric is about the:

A) variance

B) standard deviation

C) mean

D) covariance

(6) The shape of the normal curve depends on its:

A) Mean deviation

B) Quartile deviation

C) Correlation

D) Standard deviation

Answers:
1: B | 2: B| 3: A| 4: A| 5: C| 6: D

Research Team

RESEARCH TEAM 438

Keep it simple, Keep it focused

Thank you for checking our work!

Leaders:

Shuaa Khdary   Sarah AlQuwayz

Abdulrhman Alsuhaibany

Member:

Nourah Alklaib

Note Taker:

Abdulaziz Alamri   Noura Alshathri

Contact us:

Research4390@gmail.com