23

NORMAL DISTRIBUTION

KSU COLLEGE OF MEDICINE
2019 - 2020

ACKNOWLEDGMENTS

DONE BY

AFNAN ALMUSTAFA

ASEEL BADUKHON

Special thanks to SARAH ALANZI & 436 TEAM

TABLE OF **CONTENTS**

NORMAL DISTRIBUTION

LECTURE **OBJECTIVES**

By the end of this lecture, I am able to understand:

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

OVERVIEW

Introduction

- **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 ?
- The Normal or Gaussian distribution is the most important continuous probability distribution in statistics.
- The term “Gaussian” refers to ‘Carl Freidrich Gauss’ who develop this distribution.
- The word ‘**normal**’ here does not mean ‘ordinary’ or ‘common’ nor does it mean ‘**disease-free**’.
- It simply means that the distribution confirms to a certain formula and shape.

Gaussian Distribution

- Many biologic variables follow this pattern:

Hemoglobin, Cholesterol, Serum Electrolytes, Blood pressures, age, weight, height.

- One can use this information to define what is normal and what is extreme.
- In clinical medicine 95% or 2 Standard deviations around the mean is normal.
    - Clinically, 5% of “normal” individuals are labeled as extreme/abnormal.
        - We just accept this and move on.

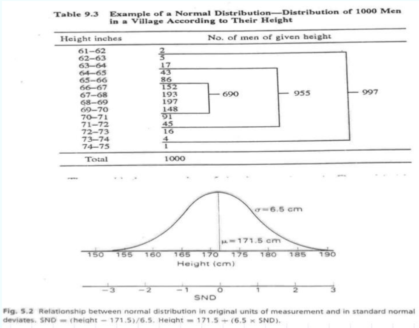
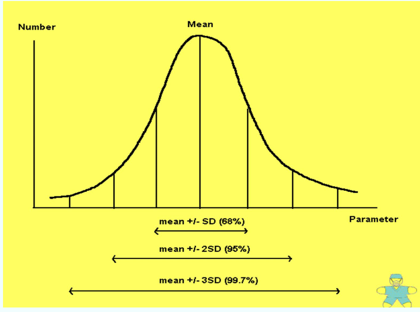


Table 9.3 Example of a Normal Distribution—Distribution of 1000 Men in a Village According to Their Height

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).



Normal Distribution

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

- Uses
  - It's application goes beyond describing distributions.
  - It is used by researchers.
  - The major use of normal distribution is the role it plays in statistical inference. Most of the statically theory based on this concept
  - It helps managers to make decisions.

- What's so Great about the Normal Distribution?
  - If you know two things (**Mean** & **Standard deviation**) you know everything about the distribution. You know the probability of any value arising.

EXAMPLE

Standardised Scores

- 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 score. How many SDs above the mean is that?

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

$$\frac{\text{My Score - Mean Score}}{\text{SD}} = \frac{100-90}{4} = 2.5$$

This is a standardised score, or z-score

Look z tables (or computer), See how often this high (or higher) score occur.

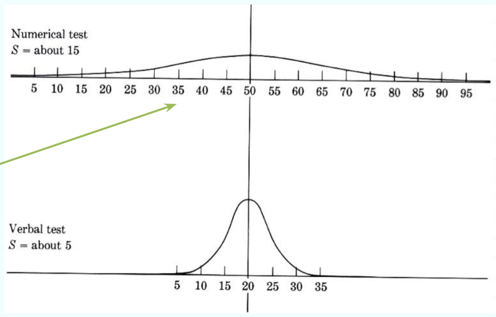
Measures of Position

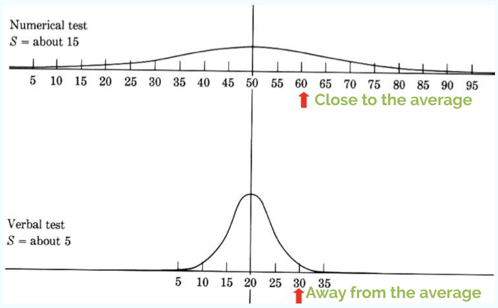
- **z Score** (or standard score)
  - The number of standard deviations that a given value x is above or below the mean. meaning that my BP is 100 and the mean is 90, the SD is 4, how many SDs do i need to reach my BP from the mean? 10 is the difference, 10 divided by 4 = 2.5 , so i need 2.5 times
  - //(90+4)=1 SD,, (90+8)= 2 SD,, (90+10)=2.5 SD
  - The Z score makes it possible, under some circumstances, to compare scores that originally had different units of measurement.

- 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 you scored a 60 on a numerical test and a 30 on a verbal test. On which test did you perform better?
  - 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? Verbal is better because you almost 2 standard deviations

Why flat?
Standard Deviation is 15 that the reason .





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

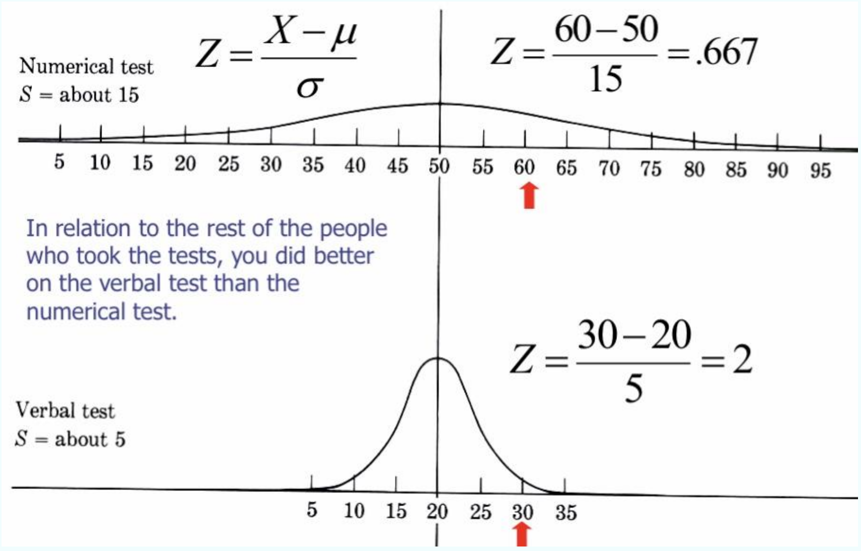
Population:

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

Sample:

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

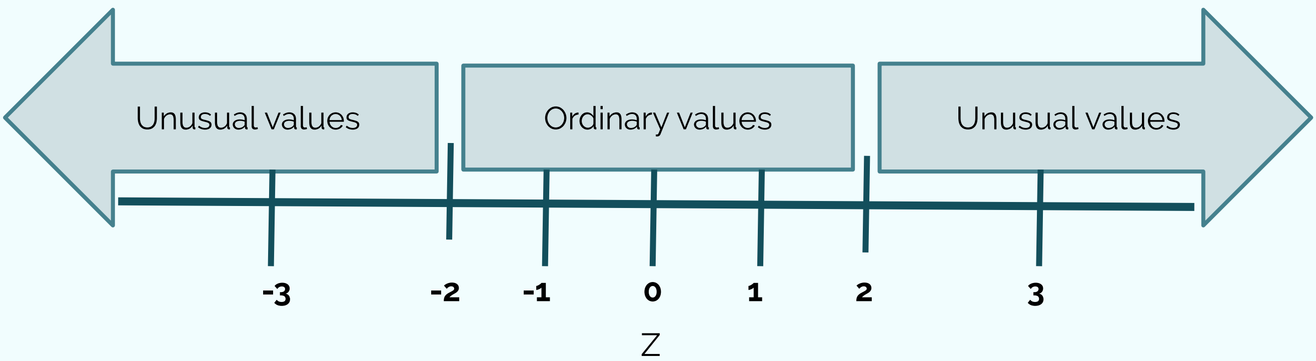
Cont'



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



INTERPRETATION

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: (Please read the heading of the table before you find the values; notice the differences)

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

Table A[a] (Continued)
PROPORTIONS OF AREA UNDER STANDARD NORMAL CURVE FOR VALUES OF z

A z	B	C
1.68	.4535	.0465
1.69	.4545	.0455
1.70	.4554	.0446
1.71	.4564	.0436
1.72	.4573	.0427
1.73	.4582	.0418
1.74	.4591	.0409
1.75	.4599	.0401
1.76	.4608	.0392
1.77	.4616	.0384
1.78	.4625	.0375
1.79	.4633	.0367
1.80	.4641	.0359
1.81	.4649	.0351
1.82	.4656	.0344
1.83	.4664	.0336
1.84	.4671	.0329
1.85	.4678	.0322
1.86	.4686	.0314
1.87	.4693	.0307
1.88	.4699	.0301
1.89	.4706	.0294
1.90	.4713	.0287
1.91	.4719	.0281
1.92	.4726	.0274
1.93	.4732	.0268
1.94	.4738	.0262
1.95	.4744	.0256
1.96	.4750	.0250
1.97	.4756	.0244
1.98	.4761	.0239
1.99	.4767	.0233
2.00	.4772	.0228
2.01	.4778	.0222
2.02	.4783	.0217
2.03	.4788	.0212
2.04	.4793	.0207
2.05	4798	.0202
2.06	.4803	.0197
2.07	.4808	.0192
2.08	.4812	.0188
2.09	.4817	.0183
2.10	.4821	.0179
2.11	.4826	.0174
2.12	.4830	.0170
2.13	.4834	.0166
2.14	.4838	.0162
2.15	.4842	.0158
2.16	.4846	.0154
2.17	.4850	.0150
2.18	.4854	.0146
2.19	.4857	.0143
2.20	.4861	.0139
2.21	.4864	.0136
2.22	.4868	.0132
2.23	.4871	.0129
2.24	.4875	.0125
2.25	.4878	.0122
2.26	.4881	.0119
2.27	.4884	.0116
2.28	.4887	.0113
2.29	.4890	.0110
2.30	.4893	.0107
2.31	.4896	.0104
2.32	.4898	.0102
2.33	.4901	.0099
2.34	.4904	.0096
2.35	.4906	.0094
2.36	.4909	.0091
2.37	.4911	.0089
2.38	.4913	.0087
2.39	.4916	.0084
2.40	.4918	.0082
2.41	.4920	.0080
2.42	.4922	.0078
2.43	.4925	.0075
2.44	.4927	.0073
2.45	.4929	.0071
2.46	.4931	.0069
2.47	.4932	.0068
2.48	.4934	.0066
2.49	.4936	.0064
2.50	.4938	.0062
2.51	.4940	.0060
2.52	.4941	.0059
2.53	.4943	.0057
2.54	.4945	.0055
2.55	.4946	.0054
2.56	.4948	.0052
2.57	.4949	.0051
2.58	.4951	.0049
2.59	.4952	.0048
2.60	.4953	.0047
2.61	.4955	.0045
2.62	.4956	.0044
2.63	.4957	.0043
2.64	.4959	.0041
2.65	.4960	.0040
2.66	.4961	.0039
2.67	.4962	.0038
2.68	.4963	.0037
2.69	.4964	.0036
2.70	.4965	.0035
2.71	.4966	.0034
2.72	.4967	.0033
2.73	.4968	.0032
2.74	.4969	.0031
2.75	.4970	.0030
2.76	.4971	.0029
2.77	.4972	.0028
2.78	.4973	.0027
2.79	.4974	.0026
2.80	.4974	.0026
2.81	.4975	.0025
2.82	.4976	.0024
2.83	.4977	.0023
2.84	.4977	.0023
2.85	.4978	.0022
2.86	.4979	.0021
2.87	.4979	.0021
2.88	.4980	.0020
2.89	.4981	.0019
2.90	.4981	.0019
2.91	.4982	.0018
2.92	.4982	.0018
2.93	.4983	.0017
2.94	.4984	.0016
2.95	.4984	.0016
2.96	.4985	.0015
2.97	.4985	.0015
2.98	.4986	.0014
2.99	.4986	.0014
3.00	.4987	.0013
3.01	.4987	.0013
3.02	.4987	.0013
3.03	.4988	.0012
3.04	.4988	.0012
3.05	.4989	.0011
3.06	.4989	.0011
3.07	.4989	.0011
3.08	.4990	.0010
3.09	.4990	.0010
3.10	.4990	.0010
3.11	.4991	.0009
3.12	.4991	.0009
3.13	.4991	.0009
3.14	.4992	.0008
3.15	.4992	.0008
3.16	.4992	.0008
3.17	.4992	.0008
3.18	.4993	.0007
3.19	.4993	.0007
3.20	.4993	.0007
3.21	.4993	.0007
3.22	.4994	.0006
3.23	.4994	.0006
3.24	.4994	.0006
3.25	.4994	.0006
3.30	.4995	.0005
3.35	.4996	.0004
3.40	.4997	.0003
3.45	.4997	.0003
3.50	.4998	.0002
3.60	.4998	.0002
3.70	.4999	.0001
3.80	.4999	.0001
3.90	.49995	.00005
4.00	.49997	.00003
−z A′	B′	C′

Reading the Z table: (Pay attention to the color code)

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

INTERPRETATION

Z scores & the Normal Distribution

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.

EXAMPLE

Assuming the normal heart rate (H.R) in normal healthy individuals is normally distributed with:

Mean = 70 and Standard Deviation =10 beats/min

Then: (next slide for the graphs and dr's notes)

1) What area under the curve is above 80 beats/min?

Ans: 0.16 (16%)

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%)

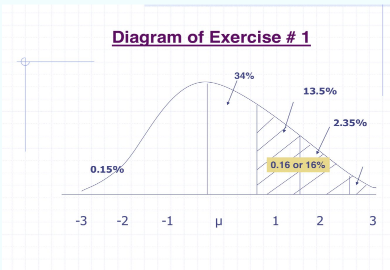
4) What area of the curve is above 100 beats/min?

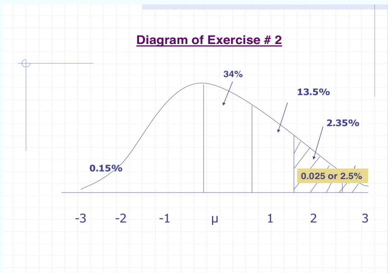
Ans: 0.0015 (0.15%)

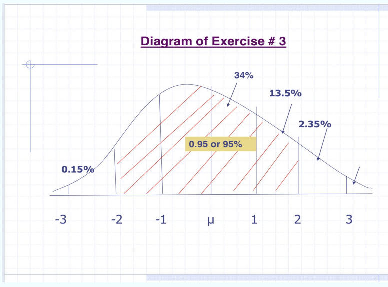
5) What area of the curve is below 40 beats per min or above 100 beats per min?

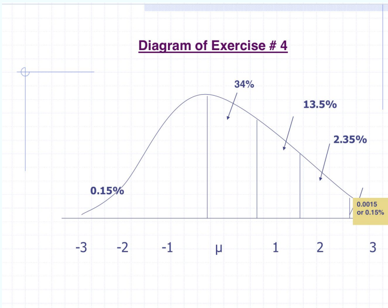
Ans: 0.0015 for each tail or 0.3%

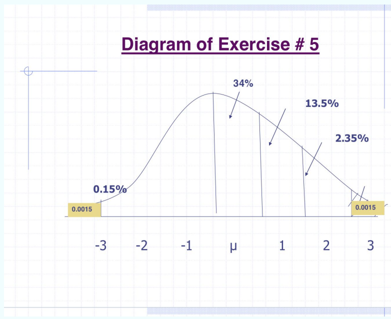
EXAMPLE











1-**How much percentage of patients are their heart rate is 80 b\m and above?** Total are 100%.//1 standard deviation cover 68%, half of 68%(we did this step because the SD covers 10 above and below 70 and in this case we only want above so we take the half) ,34% Subtract from 50 = 16%(we did this step because 34% are above the mean (70) with a range of 10(70+10=80) but what about those above it?(and our case we only want above 80) we subtract from 50 because this 50% represents those above the mean and 34% are within (70--80) and we want above 80 which leaves us with the other part of the 50% (50-34=16%))..who have a beat above 80 beats/min

2-=2.5% Because we are asking 2 standard deviations. 2 standard deviations (90) , mean is 70 , Probability .025 or 2.5%

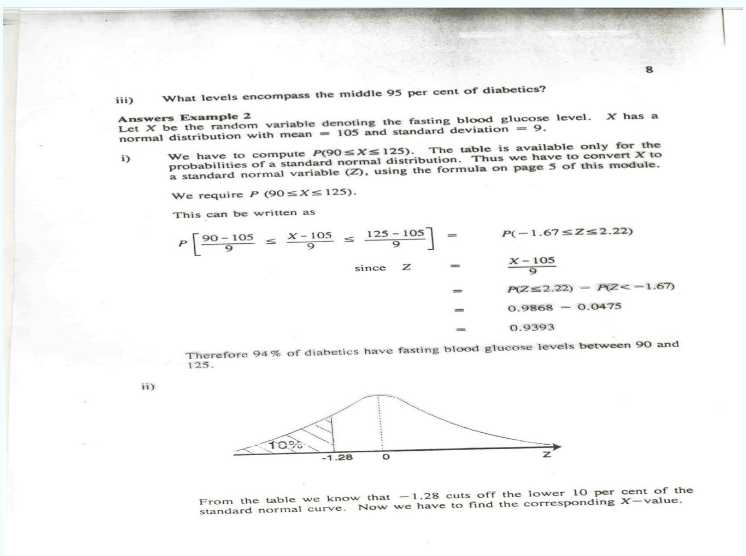
3-0.95 or 95% because of 2SD.

4-0.0015.(3 standard deviations). so small area in the extreme right side.

5-Extreme 3 standard deviations - extreme +3 standard deviations on upper side =because it crosses 3 standard Deviations

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 ?

We have to get 2 z score, it is not a whole number to apply the rule it is in fraction.
90 - 105 \ 9 = -1.6 this is one z score
125 - 105 \ 9 = 2.2 this is the second z score. then you continue solving as in the picture below.


8

iii) What levels encompass the middle 95 per cent of diabetics?

Answers Example 2
Let X be the random variable denoting the fasting blood glucose level. X has a normal distribution with mean = 105 and standard deviation = 9.

i) We have to compute $P(90 \le X \le 125)$. 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 page 5 of this module.

We require $P\ (90 \le X \le 125)$.

This can be written as

$$P\left[\frac{90-105}{9} \le \frac{X-105}{9} \le \frac{125-105}{9}\right] = P(-1.67 \le Z \le 2.22)$$

$$\text{since} \quad Z = \frac{X-105}{9}$$

$$= P(Z \le 2.22) - P(Z < -1.67)$$

$$= 0.9868 - 0.0475$$

$$= 0.9393$$

Therefore 94% of diabetics have fasting blood glucose levels between 90 and 125.

ii)

10% -1.28 0 Z

From the table we know that -1.28 cuts off the lower 10 per cent of the standard normal curve. Now we have to find the corresponding X-value.