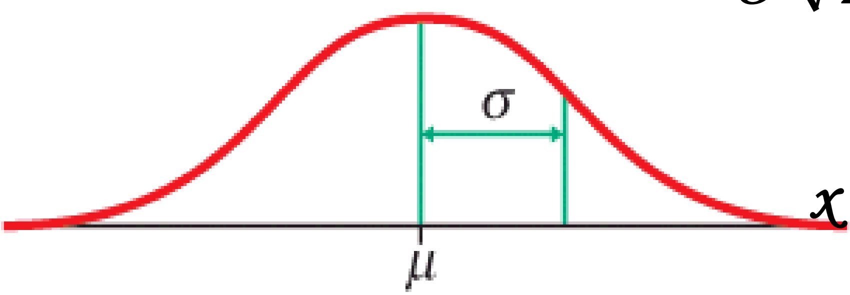
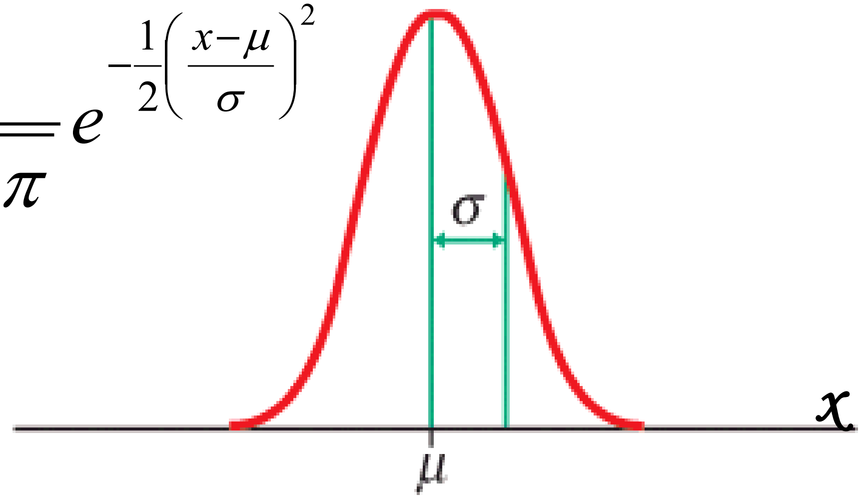
Normal Distribution

Normal – or Gaussian – distribution is a family of symmetrical, bell-shaped density curves defined by a mean μ (*mu*) and a standard deviation σ (*sigma*) : *N*(μ,σ).

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





$$e = 2.71828\ldots \text{ The base of the natural logarithm}$$

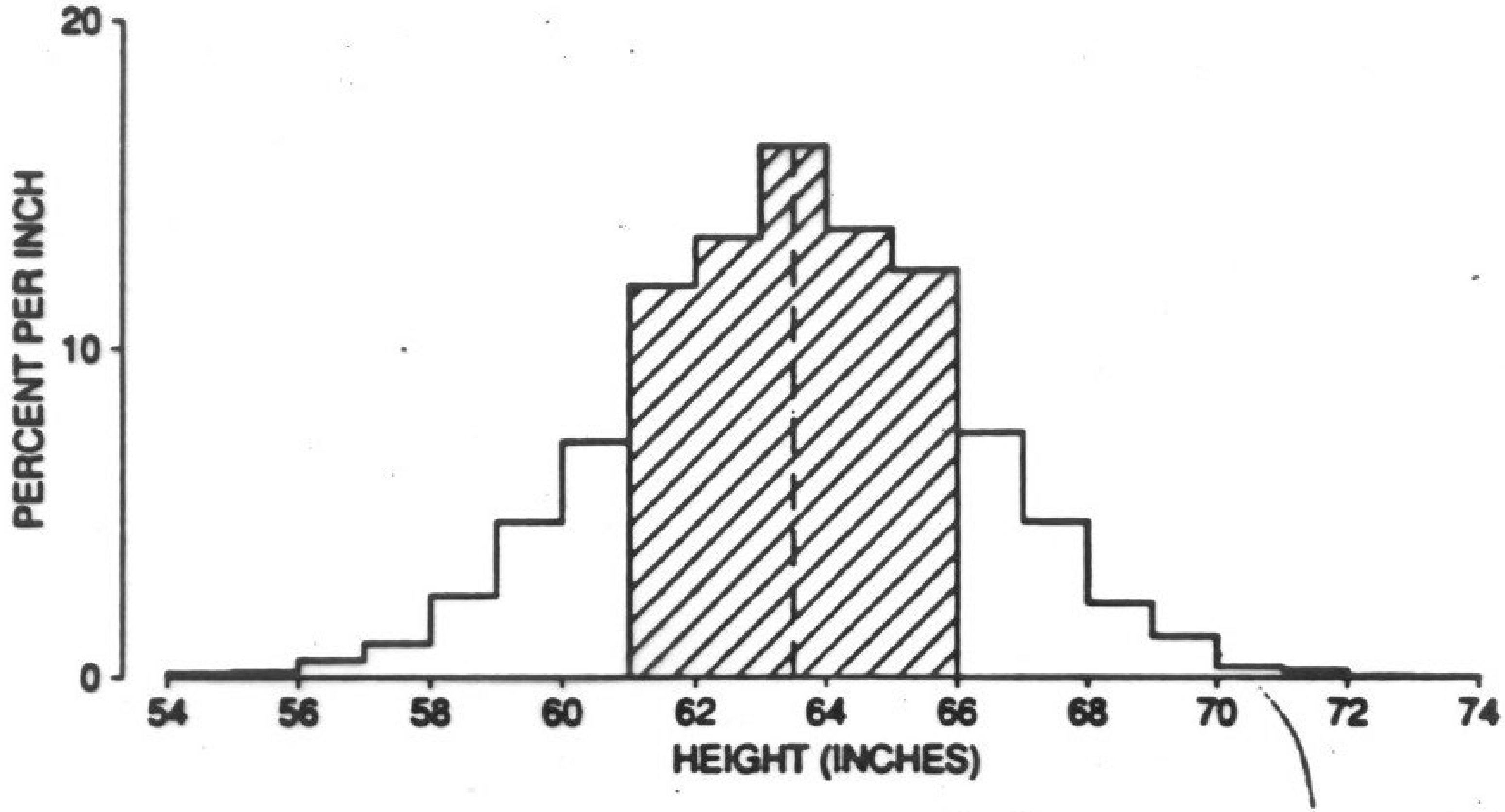
$$\pi = pi = 3.14159\ldots$$

Height of American Women

Figure 8. The SD and the histogram for the heights of the 6,588 women age 18–74 in the HANES sample. The average of 63.5 inches is marked by a dashed vertical line. The region within one SD of the average is shaded: 67% of the women differed from average by one SD (2.5 inches) or less.

$\bar{x} = 63.5$ [inches]

$s = 2.5$ [inches]



Source: Freedman, Pisani and Purves: Statistics

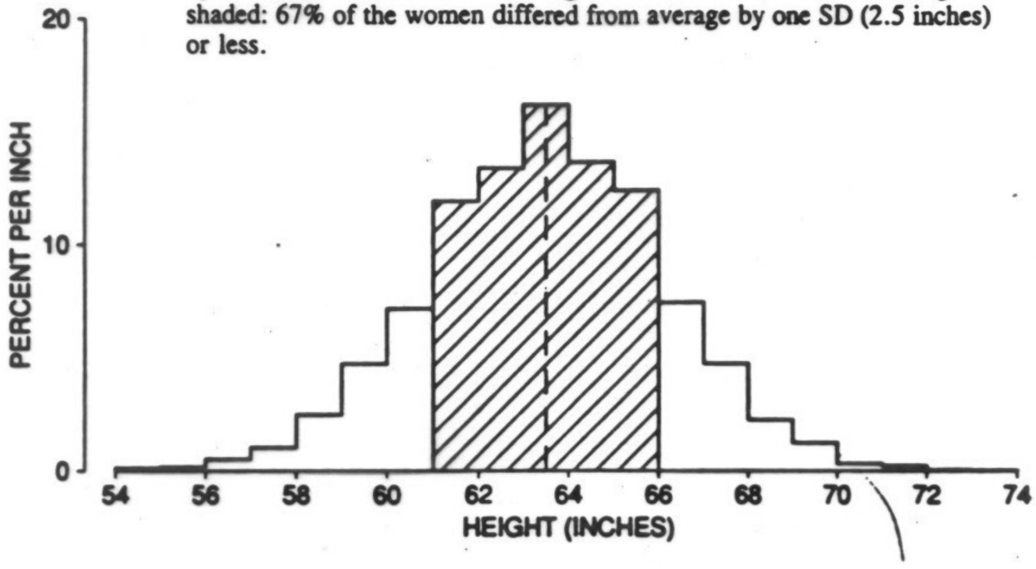
Empirical Rule

$\bar{x} \pm s$ contains approximately 68% of the sample observations

$\bar{x} \pm 2s$ contains approximately 95% of the sample observations

$\bar{x} \pm 3s$ contains approximately 99.7 % of the sample observations

Figure 8. The SD and the histogram for the heights of the 6,588 women age 18–74 in the HANES sample. The average of 63.5 inches is marked by a dashed vertical line. The region within one SD of the average is shaded: 67% of the women differed from average by one SD (2.5 inches) or less.



Source: Freedman, Pisani and Purves: Statistics

Empirical Rule

About two-thirds of the HANES women differed from the average by less than one SD.

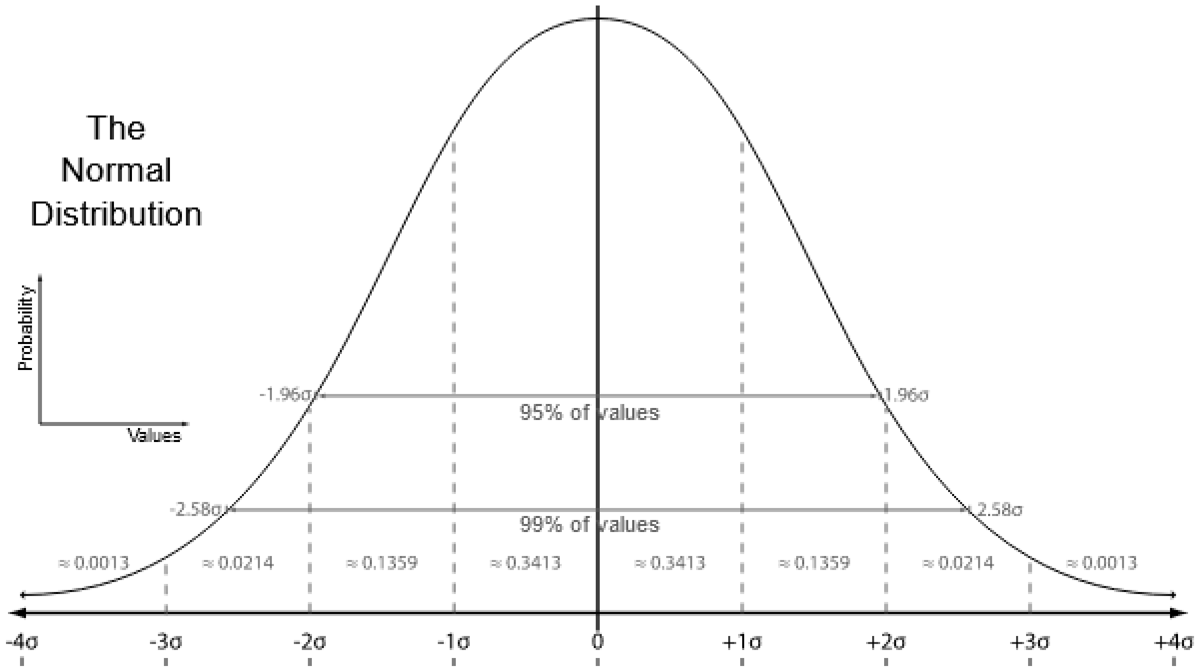
Source: Freedman, Pisani and Purves: Statistics

Normal Distribution

$\mu \pm \sigma$ contains approximately 68% of the population observations

$\mu \pm 2\sigma$ contains approximately 95% of the population observations

$\mu \pm 3\sigma$ contains approximately 99.7% of the population observations



CASE : Statistics Test Scores

- A sample of students obtained the following scores on a statistics tests.

65	13	75	89	92	73	82	85
92	87	95	62	79	82	91	87
65	81	92	86	31	63	74	85

$$\bar{x} = 76.08$$

$$s = 19.5$$

$$\bar{x} \pm s = [56.58, 95.58]$$

Actual: 22/24 = 91.67%

$$\bar{x} \pm 2s = [37.08, 115.08]$$

Actual: 22/24 = 91.67%

$$\bar{x} \pm 3s = [17.58, 134.58]$$

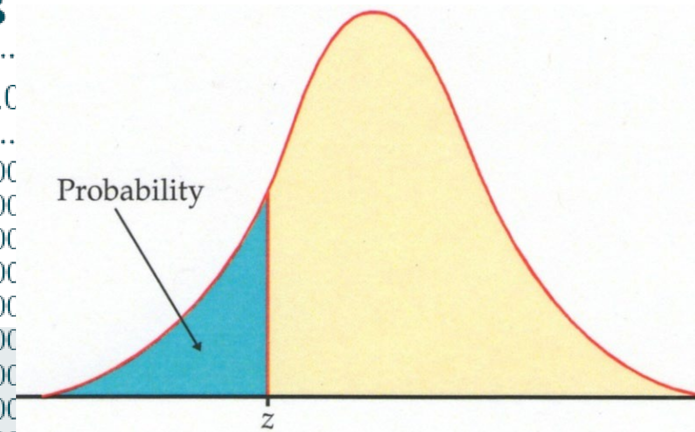
Actual: 23/24 = 95.83%

Using the standard Normal table

$P(Z \leq -2.33) = ?$

TABLE A Standard normal probabilities

z	.00	.01	.02	.03	.04	.05	.0			
−3.4	.0003	.0003	.0003	.0003	.0003	.0003	.00			
−3.3	.0005	.0005	.0005	.0004	.0004	.0004	.00			
−3.2	.0007	.0007	.0006	.0006	.0006	.0006	.00			
−3.1	.0010	.0009	.0009	.0009	.0008	.0008	.00			
−3.0	.0013	.0013	.0013	.0012	.0012	.0011	.00			
−2.9	.0019	.0018	.0018	.0017	.0016	.0016	.00			
−2.8	.0026	.0025	.0024	.0023	.0023	.0022	.00			
−2.7	.0035	.0034	.0033	.0032	.0031	.0030	.00			
−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



(...)

Standard Normal Distribution

$Z \sim N(0,1)$

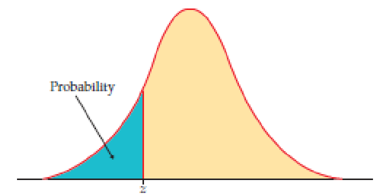
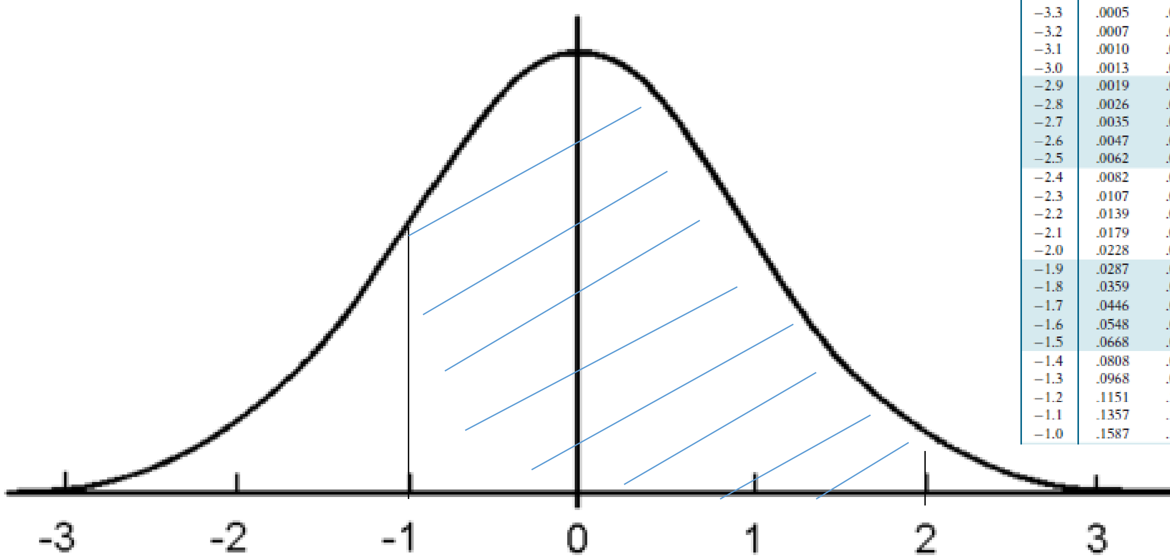


Table entry for z is the area under the standard Normal curve to the left of z.

TABLE A Standard Normal Probabilities

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

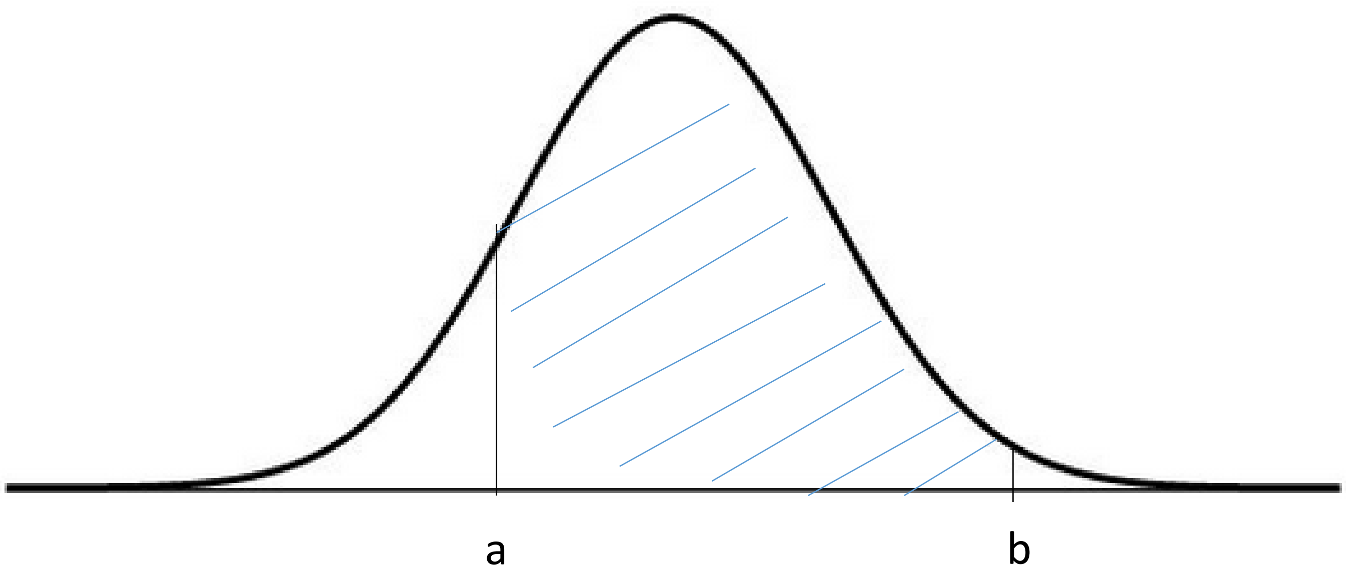


$$P(-1 \leq Z \leq 2) =$$

$$0.9772 - 0.1587 = 0.8185$$

Non-Standard Normal Distribution

$X \sim N(\mu, \sigma)$ $\quad X = \mu + \sigma Z \quad$ $Z = \dfrac{X-\mu}{\sigma} \sim N(0,1)$



$P(a \le X \le b) = P(z_a \le Z \le z_b)$

$z_a = \dfrac{a-\mu}{\sigma} \qquad z_b = \dfrac{b-\mu}{\sigma}$