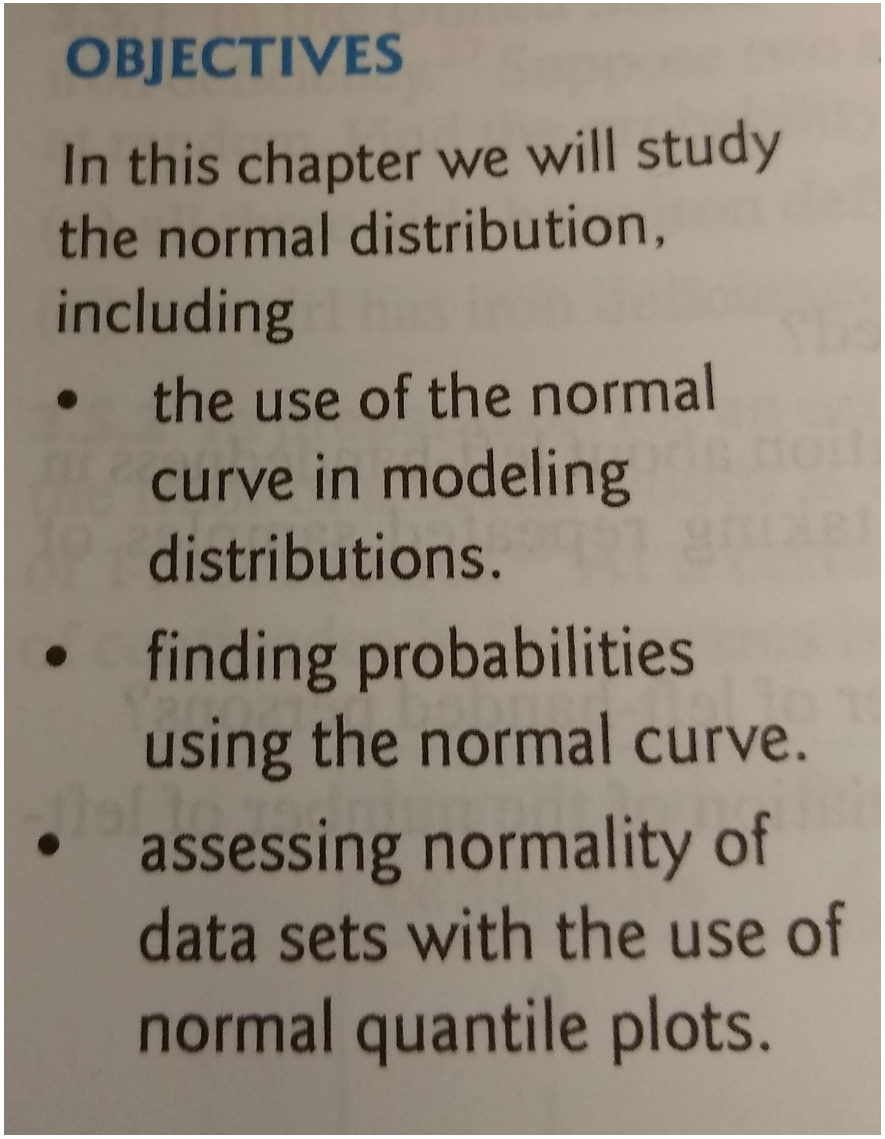
Hoofdstuk 4

De normaal verdeling

(Reference: Chapter 4 van Myra L. Samuels et al.)

Objectives



OBJECTIVES

In this chapter we will study the normal distribution, including

- the use of the normal curve in modeling distributions.
- finding probabilities using the normal curve.
- assessing normality of data sets with the use of normal quantile plots.

(Reference: Chapter 4 van Myra L. Samuels et al.)

Heeft onze studentpopulatie een normaalverdeling?

- Iedere student noemt even zijn of haar lengte
- Ik verwerk dit in een R-script om e.e.a. te visualiseren en uit te leggen

(Reference: Chapter 4 van Myra L. Samuels et al.)

Normaal verdeling

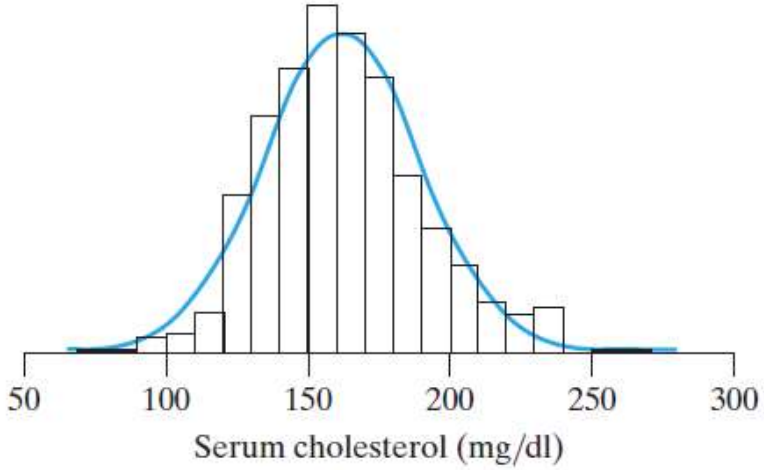


Figure 4.1.1 Distribution of serum cholesterol in 727 12- to 14-year-old children

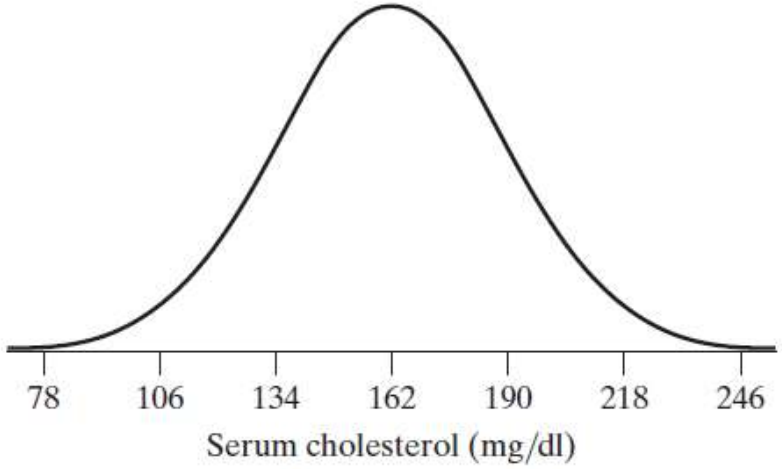


Figure 4.1.2 Normal distribution of serum cholesterol, with $\mu = 162$ mg/dl and $\sigma = 28$ mg/dl

(Reference: Chapter 4 van Myra L. Samuels et al.; 4e druk)

Meetfout

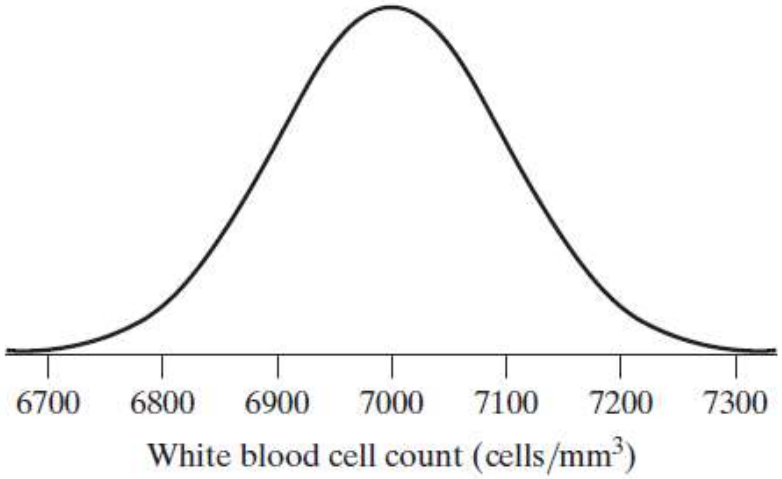


Figure 4.1.5 Normal distribution of repeated white blood cell counts of a blood specimen whose true value is $\mu = 7000$ cells/mm^3. The standard deviation is $\sigma = 100$ cells/mm^3.

(Reference: Chapter 4 van Myra L. Samuels et al.; 4e druk)

De normaal curve

If a variable Y follows a normal distribution with mean μ and standard deviation σ, then the density curve of the distribution of Y is given by the following formula:

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

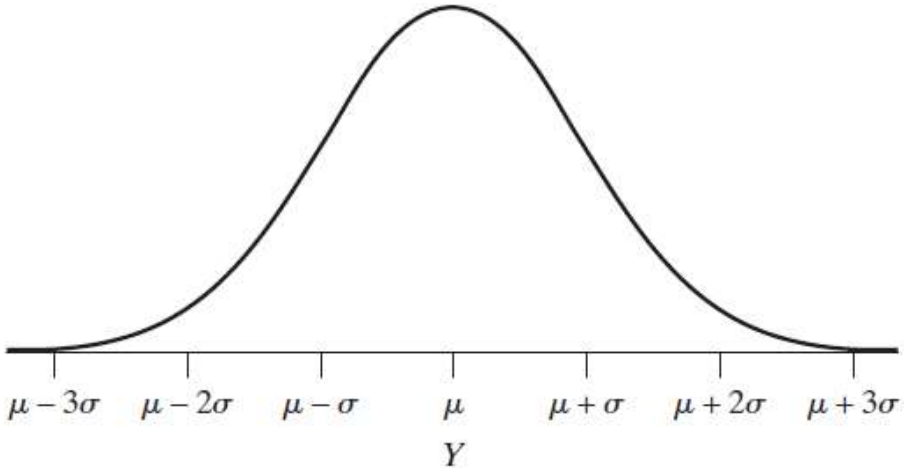
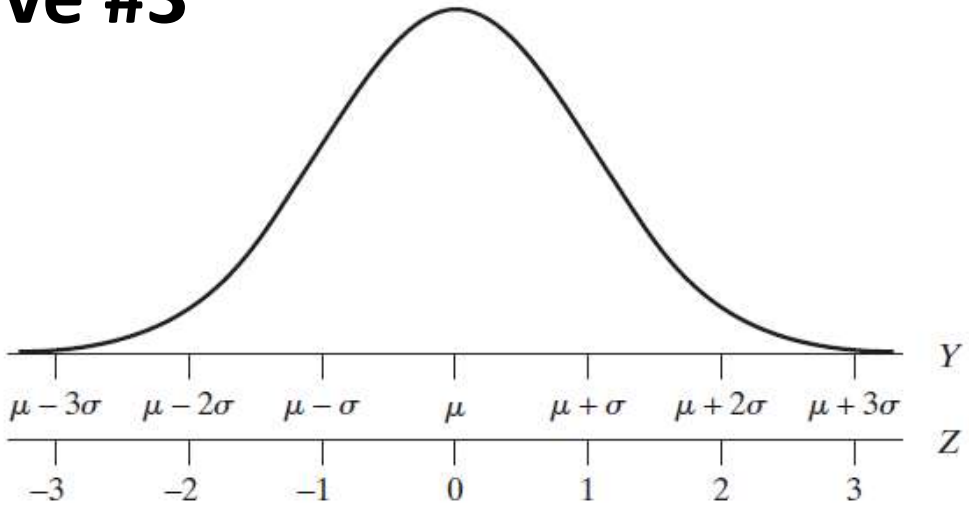


Figure 4.2.1 A normal curve with mean μ and standard deviation σ

(Reference: Chapter 4 van Myra L. Samuels et al.; 4e druk)

De normaal curve #3



As Figure 4.3.1 indicates, the Z scale measures standard deviations from the mean: $z = 1.0$ corresponds to 1.0 standard deviation above the mean; $z = -2.5$ corresponds to 2.5 standard deviations below the mean, and so on. The Z scale is referred to as a **standardized scale**.

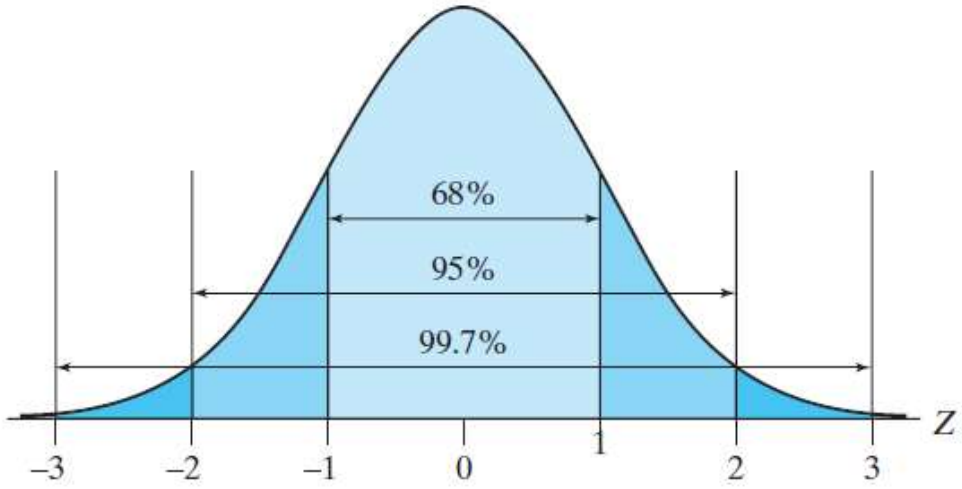
The correspondence between the Z scale and the Y scale can be expressed by the formula given in the following box.

Standardization Formula

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

(Reference: Chapter 4 van Myra L. Samuels et al.; 4e druk)

De Z-score (standaard normaalverdeling)



(Reference: Chapter 4 van Myra L. Samuels et al.; 4e druk)

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

De Z-score #2

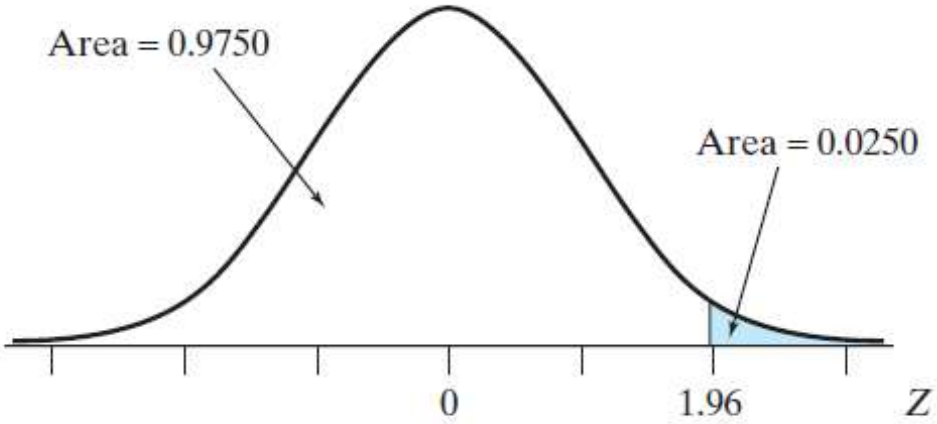


Figure 4.3.11 Area under the normal curve above 1.96

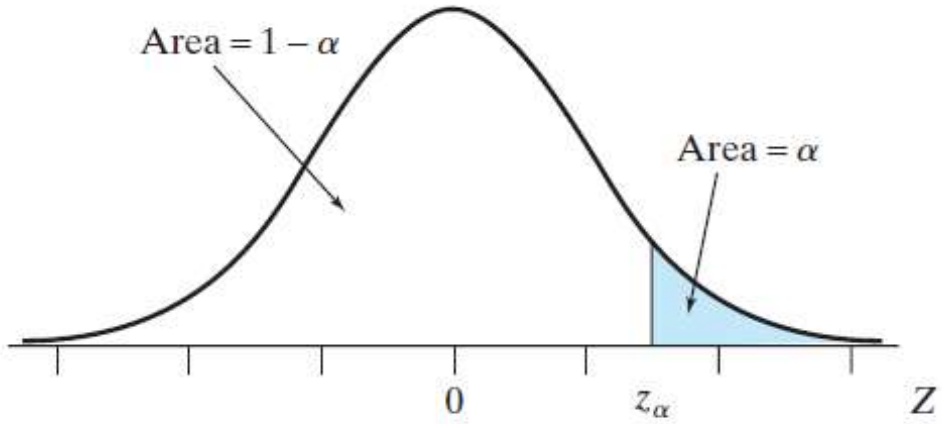


Figure 4.3.12 Area under the normal curve above α

(Reference: Chapter 4 van Myra L. Samuels et al.; 4e druk)

De oppervlakte onder de Z-curve

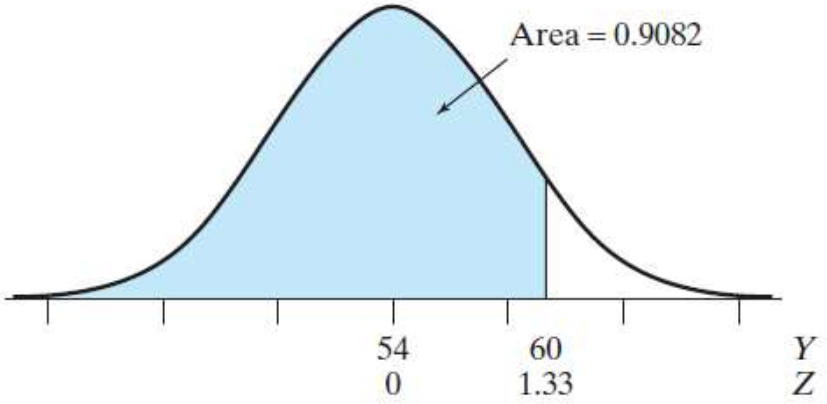
(a) What percentage of the fish are less than 60 mm long?

Figure 4.3.7 shows the population density curve, with the desired area indicated by shading. In order to use Table 3, we convert the limits of the area from the Y scale to the Z scale, as follows:

For $y = 60$, the z score is

$$z = \frac{y - \mu}{\sigma} = \frac{60 - 54}{4.5} = 1.33$$

Thus, the question "What percentage of the fish are less than 60 mm long?" is equivalent to the question "What is the area under the standard normal curve below the z value of 1.33?" Looking up $z = 1.33$ in Table 3, we find that the area is 0.9082; thus, 90.82% of the fish are less than 60 mm long.



(Reference: Chapter 4 van Myra L. Samuels et al.; 4e drul

Vragen?