Uniform, Normal and Standard Normal Distribution

From Micheal Barron Book

4.2.1 Uniform distribution

Uniform distribution plays a unique role in stochastic modeling. As we shall see in Chapter 5, a random variable with any thinkable distribution can be generated from a Uniform random variable. Many computer languages and software are equipped with a random number generator that produces Uniform random variables. Users can convert them into variables with desired distributions and use for computer simulation of various events and processes.

Also, Uniform distribution is used in any situation when a value is picked "at random" from a given interval; that is, without any preference to lower, higher, or medium values. For example, locations of errors in a program, birthdays throughout a year, and many continuous random variables modulo 1, modulo 0.1, 0.01, etc., are uniformly distributed over their corresponding intervals.

To give equal preference to all values, the Uniform distribution has a *constant* density (Figure 4.4). On the interval (a, b), its density equals

$$f(x) = \frac{1}{b-a}, \quad a < x < b,$$

because the rectangular area below the density graph must equal 1.

For the same reason, $|b - a|$ has to be a finite number. There does not exist a Uniform distribution on the entire real line. In other words, if you are asked to choose a random number from $(-\infty, +\infty)$, you cannot do it uniformly.

The Uniform property

For any $h > 0$ and $t \in [a, b-h]$, the probability

$$\boldsymbol{P}\left\{\, t < X < t+h \,\right\} = \int_t^{t+h} \frac{1}{b-a}\,dx = \frac{h}{b-a}$$

is *independent* of t. This is the *Uniform property*: the probability is only determined by the length of the interval, but not by its location.

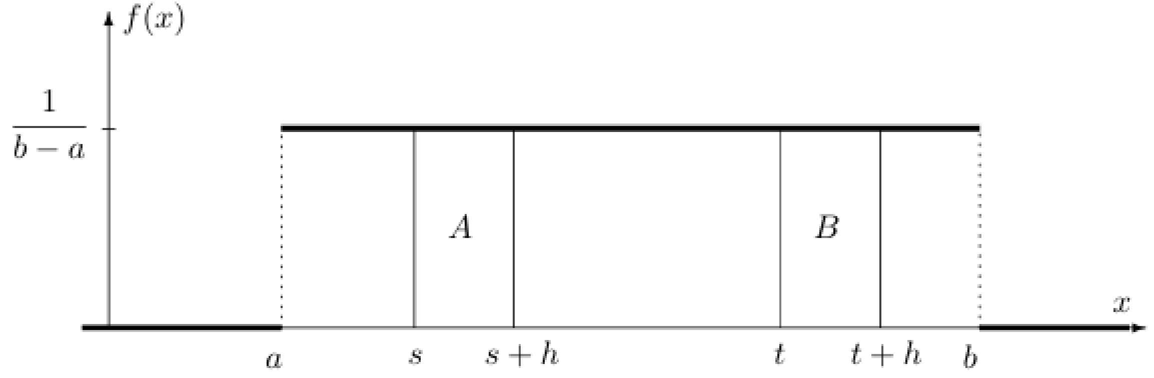


FIGURE 4.4: *The Uniform density and the Uniform property.*

Example 4.3. In Figure 4.4, rectangles A and B have the same area, showing that $\boldsymbol{P}\{s < X < s+h\} = \boldsymbol{P}\{t < X < t+h\}$. $\Diamond$

Uniform distribution

$$\begin{aligned}
(a,b) &= \text{range of values} \\
f(x) &= \frac{1}{b-a}, \quad a<x<b \\
\mathbf{E}(X) &= \frac{a+b}{2} \\
\mathrm{Var}(X) &= \frac{(b-a)^2}{12}
\end{aligned}$$

4.2.4 Normal distribution

Normal distribution plays a vital role in Probability and Statistics, mostly because of the Central Limit Theorem, according to which sums and averages often have approximately Normal distribution. Due to this fact, various fluctuations and measurement errors that consist of accumulated number of small terms appear normally distributed.

Remark: As said by a French mathematician *Jules Henri Poincaré*, "Everyone believes in the Normal law of errors, the experimenters because they think it is a mathematical theorem, the mathematicians because they think it is an experimental fact."

Besides sums, averages, and errors, Normal distribution is often found to be a good model for physical variables like weight, height, temperature, voltage, pollution level, and for instance, household incomes or student grades.

Normal distribution has a density

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{\frac{-(x-\mu)^2}{2\sigma^2}\right\}, \qquad -\infty < x < +\infty,$$

where parameters μ and σ have a simple meaning of the expectation $\mathbf{E}(X)$ and the standard deviation $\mathrm{Std}(X)$. This density is known as the bell-shaped curve, symmetric and centered at μ, its spread being controlled by σ. As seen in Figure 4.6, changing μ shifts the curve to the left or to the right without affecting its shape, while changing σ makes it more concentrated or more flat. Often μ and σ are called *location* and *scale* parameters.

Normal distribution

$$
\begin{aligned}
\mu &= \text{expectation, location parameter} \\
\sigma &= \text{standard deviation, scale parameter} \\
f(x) &= \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{\frac{-(x-\mu)^2}{2\sigma^2}\right\}, \quad -\infty < x < \infty \\
\mathbf{E}(X) &= \mu \\
\mathrm{Var}(X) &= \sigma^2
\end{aligned}
$$

Standard Normal distribution

DEFINITION 4.3

Normal distribution with "standard parameters" $\mu = 0$ and $\sigma = 1$ is called **Standard Normal distribution**.

NOTATION

$$
\begin{aligned}
Z &= \text{Standard Normal random variable} \\
\phi(x) &= \frac{1}{\sqrt{2\pi}}\, e^{-x^2/2}, \text{ Standard Normal pdf} \\
\Phi(x) &= \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}}\, e^{-z^2/2} dz, \text{ Standard Normal cdf}
\end{aligned}
$$

A Standard Normal variable, usually denoted by Z, can be obtained from a non-standard Normal(μ, σ) random variable X by *standardizing*, that is, subtracting the mean and dividing by the standard deviation,

$$Z = \frac{X - \mu}{\sigma}. \tag{4.16}$$

Unstandardizing Z, we can reconstruct the initial variable X,

$$X = \mu + \sigma Z. \tag{4.17}$$

Using these transformations, any Normal random variable can be obtained from a Standard Normal variable Z; therefore, we need a table of Standard Normal Distribution only (Table A4).

To find $\Phi(z)$ from Table A4, we locate a row with the first two digits of z and a column with the third digit of z and read the probability $\Phi(z)$ at their intersection. Notice that $\Phi(z) \approx 0$ (is "practically" zero) for all $z < -3.9$, and $\Phi(z) \approx 1$ (is "practically" one) for all $z > 3.9$.

Example 4.10 (Computing Standard Normal probabilities). For a Standard Normal random variable Z,

$$\begin{aligned}\boldsymbol{P}\{Z<1.35\} &= \Phi(1.35)=0.9115\\ \boldsymbol{P}\{Z>1.35\} &= 1-\Phi(1.35)=0.0885\\ \boldsymbol{P}\{-0.77<Z<1.35\} &= \Phi(1.35)-\Phi(-0.77)=0.9115-0.2206=0.6909.\end{aligned}$$

according to Table A4. Notice that $\boldsymbol{P}\{Z<-1.35\}=0.0885=\boldsymbol{P}\{Z>1.35\}$, which is explained by the symmetry of the Standard Normal density in Figure 4.6. Due to this symmetry, "the left tail," or the area to the left of (-1.35) equals "the right tail," or the area to the right of 1.35. $\Diamond$

In fact, the symmetry of the Normal density, mentioned in this example, allows to obtain the first part of Table A4 on p. 432 directly from the second part,

$$\Phi(-z)=1-\Phi(z) \qquad \text{for } -\infty<z<+\infty.$$

To compute probabilities about an arbitrary Normal random variable X, we have to standardize it first, as in (4.16), then use Table A4.

Example 4.11 (COMPUTING NON-STANDARD NORMAL PROBABILITIES). Suppose that the average household income in some country is 900 coins, and the standard deviation is 200 coins. Assuming the Normal distribution of incomes, compute the proportion of "the middle class," whose income is between 600 and 1200 coins.

Solution. Standardize and use Table A4. For a Normal($\mu = 900$, $\sigma = 200$) variable X,

$$
\begin{aligned}
\boldsymbol{P}\{600 < X < 1200\} &= \boldsymbol{P}\left\{\frac{600-\mu}{\sigma} < \frac{X-\mu}{\sigma} < \frac{1200-\mu}{\sigma}\right\} \\
&= \boldsymbol{P}\left\{\frac{600-900}{200} < Z < \frac{1200-900}{200}\right\} = \boldsymbol{P}\{-1.5 < Z < 1.5\} \\
&= \Phi(1.5) - \Phi(-1.5) = 0.9332 - 0.0668 = \underline{0.8664}.
\end{aligned}
$$

$\Diamond$

So far, we were computing probabilities of clearly defined events. These are *direct* problems. A number of applications require solution of an *inverse problem*, that is, finding a value of x given the corresponding probability.

Example 4.12 (Inverse problem). The government of the country in Example 4.11 decides to issue food stamps to the poorest 3% of households. Below what income will families receive food stamps?

Solution. We need to find such income x that $\boldsymbol{P}\{X < x\} = 3\% = 0.03$. This is an equation that can be solved in terms of x. Again, we standardize first, then use the table:

$$\boldsymbol{P}\{X < x\} = \boldsymbol{P}\left\{Z < \frac{x-\mu}{\sigma}\right\} = \Phi\left(\frac{x-\mu}{\sigma}\right) = 0.03,$$

from where

$$x = \mu + \sigma\Phi^{-1}(0.03).$$

In Table A4, we have to find the probability, the *table entry* of 0.03. We see that $\Phi(-1.88) \approx 0.03$. Therefore, $\Phi^{-1}(0.03) = -1.88$, and

$$x = \mu + \sigma(-1.88) = 900 + (200)(-1.88) = \underline{524} \text{ (coins)}$$

is the answer. In the literature, the value $\Phi^{-1}(\alpha)$ is often denoted by $z_{1-\alpha}$. $\Diamond$

Normal approximation to Binomial distribution

Binomial variables represent a special case of $S_n = X_1 + \ldots + X_n$, where all X_i have Bernoulli distribution with some parameter p. We know from Section 3.4.5 that small p allows to approximate Binomial distribution with Poisson, and large p allows such an approximation for the number of failures. For the moderate values of p (say, $0.05 \leq p \leq 0.95$) and for large n, we can use Theorem 1:

$$\text{Binomial}(n, p) \approx Normal\left(\mu = np, \sigma = \sqrt{np(1-p)}\right) \tag{4.19}$$

Continuity correction

This correction is needed when we approximate a discrete distribution (Binomial in this case) by a continuous distribution (Normal). Recall that the probability $\boldsymbol{P}\{X = x\}$ may be positive if X is discrete, whereas it is always 0 for continuous X. Thus, a direct use of (4.19) will always approximate this probability by 0. It is obviously a poor approximation.

This is resolved by introducing a *continuity correction*. Expand the interval by 0.5 units in each direction, then use the Normal approximation. Notice that

$$P_X(x) = \boldsymbol{P}\{X = x\} = \boldsymbol{P}\{x - 0.5 < X < x + 0.5\}$$

is true for a Binomial variable X; therefore, the continuity correction does not change the event and preserves its probability. It makes a difference for the Normal distribution, so every time when we approximate some discrete distribution with some continuous distribution, we should be using a continuity correction. Now it is the probability of an interval instead of one number, and it is not zero.

Example 4.15. A new computer virus attacks a folder consisting of 200 files. Each file gets damaged with probability 0.2 independently of other files. What is the probability that fewer than 50 files get damaged?

Solution. The number X of damaged files has Binomial distribution with $n = 200$, $p = 0.2$, $\mu = np = 40$, and $\sigma = \sqrt{np(1-p)} = 5.657$. Applying the Central Limit Theorem with the continuity correction,

$$\begin{aligned}\boldsymbol{P}\{X < 50\} &= \boldsymbol{P}\{X < 49.5\} = \boldsymbol{P}\left\{\frac{X-40}{5.657} < \frac{49.5-40}{5.657}\right\} \\ &= \Phi(1.68) = \underline{0.9535}.\end{aligned}$$

Notice that the properly applied continuity correction replaces 50 with 49.5, not 50.5. Indeed, we are interested in the event that X is *strictly* less than 50. This includes all values up to 49 and corresponds to the interval $[0, 49]$ that we *expand* to $[0, 49.5]$. In other words, events $\{X < 50\}$ and $\{X < 49.5\}$ are the same; they include the same possible values of X. Events $\{X < 50\}$ and $\{X < 50.5\}$ are different because the former includes $X = 50$, and the latter does not. Replacing $\{X < 50\}$ with $\{X < 50.5\}$ would have changed its probability and would have given a wrong answer. $\Diamond$

Exercises

4.16. Let Z be a Standard Normal random variable. Compute

(a) $P(Z < 1.25)$ (b) $P(Z \leq 1.25)$ (c) $P(Z > 1.25)$
(d) $P(|Z| \leq 1.25)$ (e) $P(Z < 6.0)$ (f) $P(Z > 6.0)$
(g) With probability 0.8, variable Z does not exceed what value?

4.17. For a Standard Normal random variable Z, compute

(a) $P(Z \geq 0.99)$ (b) $P(Z \leq -0.99)$ (c) $P(Z < 0.99)$
(d) $P(|Z| > 0.99)$ (e) $P(Z < 10.0)$ (f) $P(Z > 10.0)$
(g) With probability 0.9, variable Z is less than what?

4.18. For a Normal random variable X with $\mathbf{E}(X) = -3$ and $\mathrm{Var}(X) = 4$, compute

(a) $P(X \leq 2.39)$ (b) $P(Z \geq -2.39)$ (c) $P(|X| \geq 2.39)$
(d) $P(|X + 3| \geq 2.39)$ (e) $P(X < 5)$ (f) $P(|X| < 5)$
(g) With probability 0.33, variable X exceeds what value?

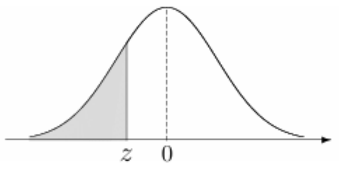
4.21. The average height of professional basketball players is around 6 feet 7 inches, and the standard deviation is 3.89 inches. Assuming Normal distribution of heights within this group,

(a) What percent of professional basketball players are taller than 7 feet?

(b) If your favorite player is within the tallest 20% of all players, what can his height be?

Table A4. Standard Normal distribution

$$\Phi(z) = \boldsymbol{P}\left\{Z \leq z\right\} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-x^2/2} dx$$

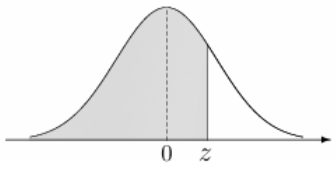


z	-0.09	-0.08	-0.07	-0.06	-0.05	-0.04	-0.03	-0.02	-0.01	-0.00
-(3.9+)	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
-3.8	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001
-3.7	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001
-3.6	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0002	.0002
-3.5	.0002	.0002	.0002	.0002	.0002	.0002	.0002	.0002	.0002	.0002
-3.4	.0002	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003
-3.3	.0003	.0004	.0004	.0004	.0004	.0004	.0004	.0005	.0005	.0005
-3.2	.0005	.0005	.0005	.0006	.0006	.0006	.0006	.0006	.0007	.0007
-3.1	.0007	.0007	.0008	.0008	.0008	.0008	.0009	.0009	.0009	.0010
-3.0	.0010	.0010	.0011	.0011	.0011	.0012	.0012	.0013	.0013	.0013
-2.9	.0014	.0014	.0015	.0015	.0016	.0016	.0017	.0018	.0018	.0019
-2.8	.0019	.0020	.0021	.0021	.0022	.0023	.0023	.0024	.0025	.0026
-2.7	.0026	.0027	.0028	.0029	.0030	.0031	.0032	.0033	.0034	.0035
-2.6	.0036	.0037	.0038	.0039	.0040	.0041	.0043	.0044	.0045	.0047
-2.5	.0048	.0049	.0051	.0052	.0054	.0055	.0057	.0059	.0060	.0062
-2.4	.0064	.0066	.0068	.0069	.0071	.0073	.0075	.0078	.0080	.0082
-2.3	.0084	.0087	.0089	.0091	.0094	.0096	.0099	.0102	.0104	.0107
-2.2	.0110	.0113	.0116	.0119	.0122	.0125	.0129	.0132	.0136	.0139
-2.1	.0143	.0146	.0150	.0154	.0158	.0162	.0166	.0170	.0174	.0179
-2.0	.0183	.0188	.0192	.0197	.0202	.0207	.0212	.0217	.0222	.0228
-1.9	.0233	.0239	.0244	.0250	.0256	.0262	.0268	.0274	.0281	.0287
-1.8	.0294	.0301	.0307	.0314	.0322	.0329	.0336	.0344	.0351	.0359
-1.7	.0367	.0375	.0384	.0392	.0401	.0409	.0418	.0427	.0436	.0446
-1.6	.0455	.0465	.0475	.0485	.0495	.0505	.0516	.0526	.0537	.0548
-1.5	.0559	.0571	.0582	.0594	.0606	.0618	.0630	.0643	.0655	.0668
-1.4	.0681	.0694	.0708	.0721	.0735	.0749	.0764	.0778	.0793	.0808
-1.3	.0823	.0838	.0853	.0869	.0885	.0901	.0918	.0934	.0951	.0968
-1.2	.0985	.1003	.1020	.1038	.1056	.1075	.1093	.1112	.1131	.1151
-1.1	.1170	.1190	.1210	.1230	.1251	.1271	.1292	.1314	.1335	.1357
-1.0	.1379	.1401	.1423	.1446	.1469	.1492	.1515	.1539	.1562	.1587
-0.9	.1611	.1635	.1660	.1685	.1711	.1736	.1762	.1788	.1814	.1841
-0.8	.1867	.1894	.1922	.1949	.1977	.2005	.2033	.2061	.2090	.2119
-0.7	.2148	.2177	.2206	.2236	.2266	.2296	.2327	.2358	.2389	.2420
-0.6	.2451	.2483	.2514	.2546	.2578	.2611	.2643	.2676	.2709	.2743
-0.5	.2776	.2810	.2843	.2877	.2912	.2946	.2981	.3015	.3050	.3085
-0.4	.3121	.3156	.3192	.3228	.3264	.3300	.3336	.3372	.3409	.3446
-0.3	.3483	.3520	.3557	.3594	.3632	.3669	.3707	.3745	.3783	.3821
-0.2	.3859	.3897	.3936	.3974	.4013	.4052	.4090	.4129	.4168	.4207
-0.1	.4247	.4286	.4325	.4364	.4404	.4443	.4483	.4522	.4562	.4602
-0.0	.4641	.4681	.4721	.4761	.4801	.4840	.4880	.4920	.4960	.5000

Table A4, continued.

Standard Normal distribution

$$\Phi(z) = \boldsymbol{P}\left\{Z \leq z\right\} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-x^2/2} dx$$



z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
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
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986
3.0	.9987	.9987	.9987	.9988	.9988	.9989	.9989	.9989	.9990	.9990
3.1	.9990	.9991	.9991	.9991	.9992	.9992	.9992	.9992	.9993	.9993
3.2	.9993	.9993	.9994	.9994	.9994	.9994	.9994	.9995	.9995	.9995
3.3	.9995	.9995	.9995	.9996	.9996	.9996	.9996	.9996	.9996	.9997
3.4	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9998
3.5	.9998	.9998	.9998	.9998	.9998	.9998	.9998	.9998	.9998	.9998
3.6	.9998	.9998	.9999	.9999	.9999	.9999	.9999	.9999	.9999	.9999
3.7	.9999	.9999	.9999	.9999	.9999	.9999	.9999	.9999	.9999	.9999
3.8	.9999	.9999	.9999	.9999	.9999	.9999	.9999	.9999	.9999	.9999
3.9+	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00