1 Some Continuous Probability Distributions

1.1 Continuous Uniform Distribution

Uniform distribution (Rectangular distribution): The density function of the continuous uniform random variable X on the interval $[A, B]$ is

$$f(x; A, B) = \begin{cases} \frac{1}{B-A}, & A \leq x \leq B \\ 0, & elsewhere \end{cases}$$

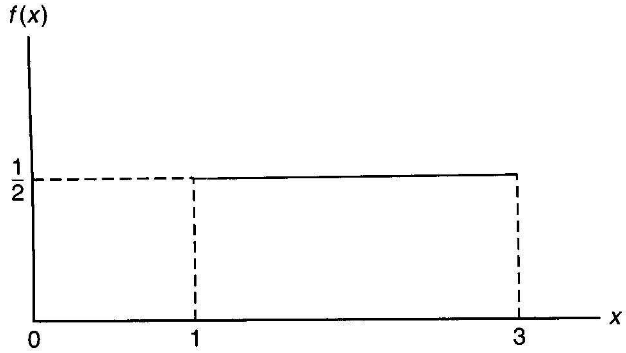

Figure 1: The density function for a random variable on the interval $[1, 3]$.

Example. Let T be the waiting for a bus when a bus comes every 30 min,

$$f(t) = \frac{1}{30}, \quad 0 \leq t \leq 30$$

- **Example 6.1**: It is assumed that length X of a conference has a uniform distribution on the interval $[0, 4]$.

  i What is the probability that team A will win the series in six games?

  $$f(x) = \begin{cases} \frac{1}{4}, & 0 \leq x \leq 4 \\ 0, & elsewhere \end{cases}$$

  ii What is the probability density function?

  $$P(X \geq 3) = \int_3^4 \frac{1}{4} dx = \frac{1}{4}$$

- **Theorem 6.1**:

  The mean and variance of the uniform distribution are

  $$\mu = \frac{A+B}{2}, \ \sigma^2 = \frac{(B-A)^2}{12}$$

- Mean is at the center of the range as we would expect.

1.2 Normal Distribution

- The most important continuous probability distribution in the entire field of statistics is the **normal distribution**.
- The **normal curve** describes approximately many phenomena that occur in nature, industry and research (human height, measurement errors, stock market!, etc.).
- In 1733, Abraham DeMoivre developed the mathematical equation of the normal curve.
- The normal distribution is often referred to as the **Gaussian distribution**, in honour of Karl Friedrich Gauss (1777-1855), who also derived its equation from a study of errors in repeated measurements of the same quantity.
- The term normal distribution is a historical accident because there is nothing particularly normal about the normal distribution and nor is there anything abnormal about other distribution.
- Normal Distribution:

  The density function of the normal random variable X, with mean μ and variance σ^2 is

  $$n(x;\mu,\sigma) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{1}{2\sigma^2}(x-\mu)^2}, \ -\infty < x < \infty$$

  A continuous random variable X having the bell-shaped distribution of Fig. 2 is called a **normal random variable**.

- Notes:
  - $(x-\mu)^2$ is squared distance from the mean
  - $e^{-\frac{1}{2\sigma^2}(x-\mu)^2}$ get smaller as $(x-\mu)^2$ gets larger

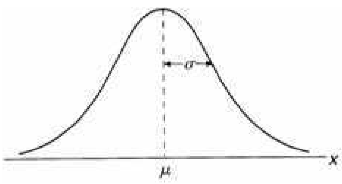

Figure 2: The normal curve.

- How fast it gets small depends on σ. Faster for small σ.
- The term $\frac{1}{\sqrt{2\pi}\sigma}$ makes sure $\int_{-\infty}^{\infty} n(x;\mu,\sigma)dx = 1$

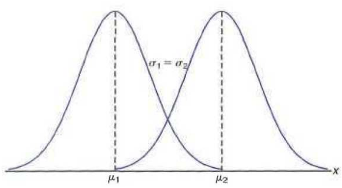

Figure 3: Normal curves with $\mu_1 < \mu_2$ and $\sigma_1 = \sigma_2$.

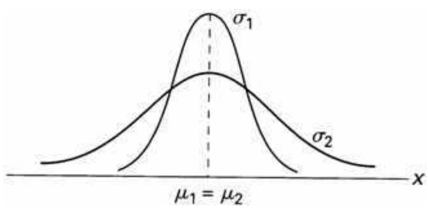

Figure 4: Normal curves with $\mu_1 = \mu_2$ and $\sigma_1 < \sigma_2$.

The properties of the normal curve

1. The mode, which is the point on the horizontal axis where the curve is a maximum, occurs at $x = \mu$.
2. The curve is symmetric about a vertical axis through the mean μ.
3. The curve has its points of inflection at $x = \mu \pm \sigma$, is concave downward if $\mu - \sigma < X < \mu + \sigma$, and is concave upward otherwise.
4. The normal curve approaches the horizontal axis asymptotically as we proceed in either direction away from the mean.

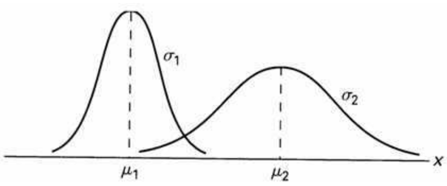

Figure 5: Normal curves with $\mu_1 < \mu_2$ and $\sigma_1 < \sigma_2$.

5. The total area under the curve and above the horizontal axis is equal to 1.
6. Both tails become dramatically thin beyond $\pm 3\sigma$ from the mean μ.

- A certain type of battery lasts on average3 years with a standard deviation of 0.5 years.
- Assuming battery lives are normally distributed,
- Find the probability that a given battery will last less than 2.3 years;
- Solution:
$$P(X < 2.3) = \int_{-\infty}^{2.3} \frac{1}{\sqrt{2\pi 0.5}} e^{-\frac{1}{2*0.5^2}(x-3)^2} dx$$
- Difficult to solve!
- Then, tabulation of normal curve areas is necessary.

1.3 Areas Under the Normal Curve

- The probability of the random variable X assuming a value between x_1 and x_2.
$$P(x_1 < X < x_2) = \int_{x_1}^{x_2} n(x; \mu, \sigma)dx = \frac{1}{\sqrt{2\pi\sigma}} \int_{x_1}^{x_2} e^{-\frac{1}{2}(\frac{x-\mu}{\sigma})^2} dx$$
- The area under the curve between any two ordinates must also depend on the values μ and σ.
- **Definition 6.1**:
The distribution of a normal random variable with mean 0 and variance 1 is called a **standard normal distribution**.

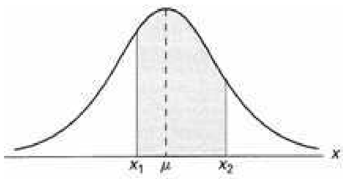

Figure 6: $P(x_1 < X < x_2)$: area of the shaded region.

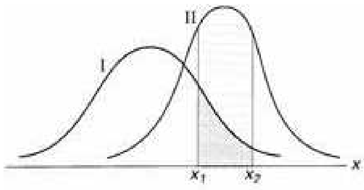

Figure 7: $P(x_1 < X < x_2)$ for different normal curves.

- Transformation: $Z = \frac{X-\mu}{\sigma}$, $z_1 = \frac{x_1-\mu}{\sigma}$, $z_2 = \frac{x_2-\mu}{\sigma}$
- Then; $E(Z) = \mu = 0$ and $\sigma_Z^2 = 1$

$$P(x_1 < X < x_2) = \frac{1}{\sqrt{2\pi}\sigma} \int_{x_1}^{x_2} e^{-\frac{1}{2}(\frac{x-\mu}{\sigma})^2} dx \Rightarrow \frac{1}{\sqrt{2\pi}} \int_{z_1}^{z_2} e^{\frac{1}{2}(z)^2} dz$$

$$= \int_{z_1}^{z_2} n(z; 0, 1) = P(z_1 < Z < z_2)$$

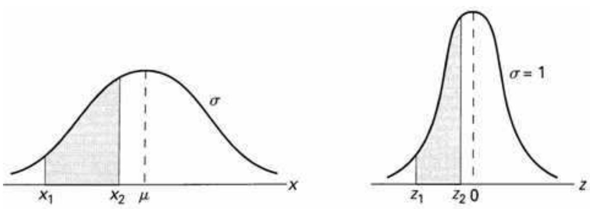

Figure 8: The original and transformed normal distributions.

Example 6.2: Given a standard normal distribution, find the area under the curve that lies

i to the right of $z = 1.84$

1 minus the area to the left of $z = 1.84$ (see Table A.3)

$$1 - 0.9671 = 0.0329$$

ii between $z = -1.97$ and $z = 0.86$

The area to the left of $z = 0.86$ minus the left of $z = -1.97$

$$0.8051 - 0.0244 = 0.7807$$

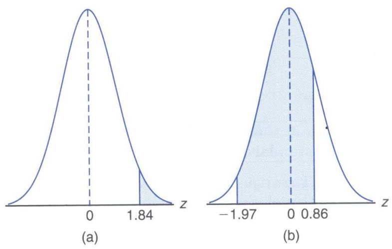

Figure 9: Areas for Example 6.2.

- Usage of the Table A.4;
  - The entries in the table are the areas under the standard normal curve between the mean, $z = 0$, and $z = X$.
  - The first column represents the values of z from 0.0 to 3.4 by increment 0.1,
  - and the first row indicates the second digit under the decimal of the corresponding values of z according to the column values.
  - Suppose we want to find the area between 0 and 1.23, then all we need to do is to read the entry where the row of 1.2 and the column of 0.03 come across.

Example 6.3: Given a standard normal distribution, find the value of k such that

i $P(Z > k) = 0.3015$

$$P(Z < k) = 1 - P(Z > k) = 1 - 0.3015 = 0.6985 \Rightarrow k = 0.52$$

TABLE A.3 (continued) Areas Under the Normal Curve

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9278	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998

ii $P(k < Z < -0.18) = 0.4197$

$$P(Z < -0.18) - P(Z < k) = 0.4286 - P(Z < k) = 0.4197 \Rightarrow k = -2.37$$

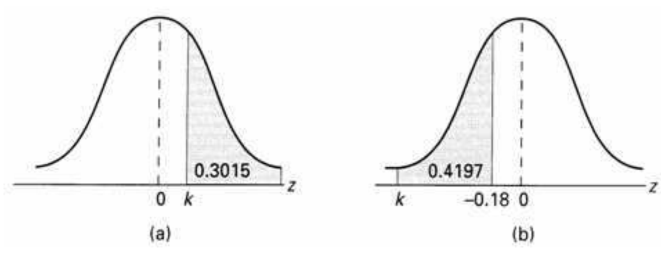

Figure 10: Areas for Example 6.3.

- **Example 6.4**: Given a random variable X having a normal distribution with $\mu = 50$ and $\sigma = 10$,

- Find the probability that X assumes a value between 45 and 62.
- Solution:

$$x_1 = 45 \; and \; x = 62 \xrightarrow{transformation} z_1 = \frac{45-50}{10} = -0.5, \; z_2 = \frac{62-50}{10} = 1.2$$

$$P(45 < X < 62) = P(-0.5 < Z < 1.2)$$
$$= P(Z < 1.2) - P(Z < -0.5) = 0.8849 - 0.3085 = 0.5764$$

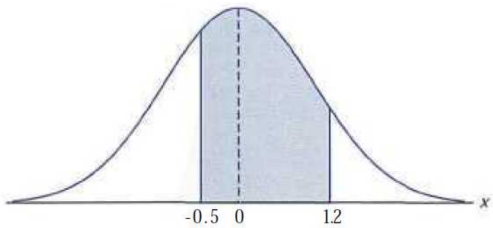

Figure 11: Area for Example 6.4.

- **Example 6.5** Given that a normal distribution with $\mu = 300$ and $\sigma = 50$, find the probability that X assumes a value greater than 362.
- Solution:

$$z = \frac{362-300}{50} 1.24$$
$$-P(X > 362) = P(Z > 1.24) = 1 - P(Z < 1.24)$$
$$= 1 - 0.8925 = 0.1075$$

- **Using the Normal Curve in Reverse**
- We might want to find the value of z corresponding to a specified probability.
- The steps:
    1. Begin with a known area or probability.

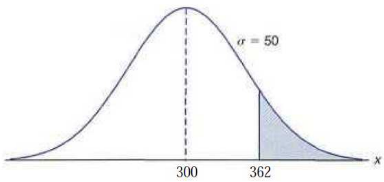

Figure 12: Area for Example 6.5.

2. Find the z values corresponding to the tabular probability that comes closest to the specified probability.
3. Determine x by rearranging the formula

$$z = \frac{x - \mu}{\sigma} \text{ to give } x = \sigma z + \mu$$

Example 6.6: Given a normal distribution with $\mu = 40$ and $\sigma = 6$, find the value of x that has

i 45% of the area to the left

From Table A.3 we find $P(Z < -0.13) = 0.45$. Hence

$$x = 6 * (-0.13) + 40 = 39.22.$$

ii 14% of the are to the right

From Table A.3, we find $P(Z < 1.08) = 086$. Hence

$$x = 6 * (1.08) + 40 = 46.48.$$

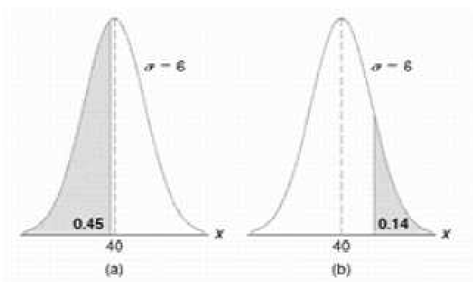

Figure 13: Areas for Example 6.6.

1.4 Applications of the Normal Distribution

- Some of the many problems for which the normal distribution is applicable are treated in the following examples.
- **Example 6.7**: A certain type of storage battery lasts, on average, 3.0 years, with a standard deviation of 0.5 year.
- Assuming that the battery lives are normally distributed, find the probability that a given battery will last less than 2.3 years.
- Solution:

$$z = \frac{2.3 - 3}{0.5} = -1.4 \Rightarrow P(X < 2.3) = P(Z < -1.4) = 0.0808$$

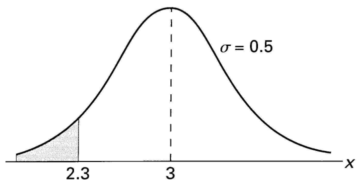

Figure 14: Area for Example 6.7.

- **Example 6.8**: An electrical firm manufactures light bulbs that have a life, before burn-out, that is normally distributed with mean equal to 800 hours and a standard deviation of 40 hours.
- Find the probability that a bulb burns between 778 and 834 hours.
- Solution:

$$z_1 = \frac{778 - 800}{40} = -0.55 \; and \; z_2 = \frac{834 - 800}{40} = 0.85$$

$$P(778 < X < 834) = P(-0.55 < Z < 0.85)$$

$$= P(Z < 0.85) - P(Z < -0.55) = 0.8023 - 0.2912 = 0.5111$$

- **Example 6.9**: The buyer sets specifications on the diameter to be 3.0 ± 0.01 cm.
- It is known that in the process the diameter of a ball bearing has a normal distribution with mean $\mu = 3.0$ and standard deviation $\sigma = 0.005$.

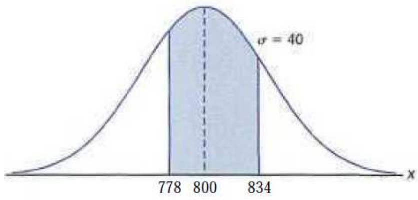

Figure 15: Area for Example 6.8.

- On the average, how many manufactured ball bearings will be scrapped?.
- Solution:

$$
\begin{aligned}
z_1 &= \frac{2.99 - 3.0}{0.005} = -2.0 \\
z_2 &= \frac{3.01 - 3.0}{0.005} = 2.0 \\
&\Rightarrow P(2.99 < X < 3.01) \\
&= P(-2.0 < Z < 2.0) \\
&= 1 - 2 * P(Z < -2.0) \\
&= 1 - 2 * 0.0228 = 0.9544
\end{aligned}
$$

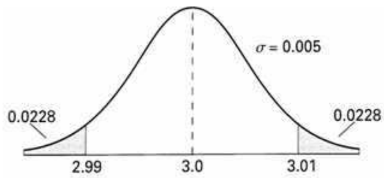

Figure 16: Area for Example 6.9.

- **Example 6.10**: Gauges are used to reject all components where a certain dimension is not within the specification $1.50 \pm d$.
- It is known that this measurement is normally distributed with mean $\mu = 1.50$ and standard deviation $\sigma = 0.2$.
- Determine the value d such that the specifications cover 95% of the measurements.

- Solution:

  From Table A.3 we know that

$$P(-1.96 < Z < 1.96) = 0.95$$

$$1.96 = \frac{(1.50 + d) - 1.50}{0.2}$$

$$\Rightarrow d = 0.2 * 1.96 = 0.392$$

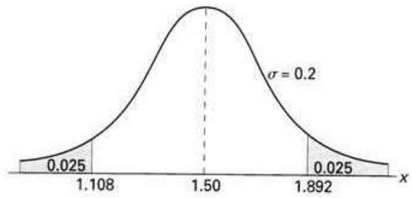

Figure 17: Specifications for Example 6.10.

- **Example 6.11**: A certain machine makes electrical resistors having a mean resistance of 40 ohms and a standard deviation of 2 ohms.
- Assuming that the resistance follows a normal distribution and can be measured to any degree of accuracy, what percentage of resistors will have a resistance exceeding 43 ohms?
- Solution:

  From Table A.3 we know that

$$z = \frac{43 - 40}{2} = 1.5$$

$$P(X > 43) = P(Z > 1.5)$$

$$= 1 - P(Z < 1.5) = 1 - 0.9332$$

$$= 0.0668$$

- **Example 6.13**: The average grade for an exam is 74, and the standard deviation is 7.
- If 12% of the class are given A's, and the grades are curved to follow a normal distribution,

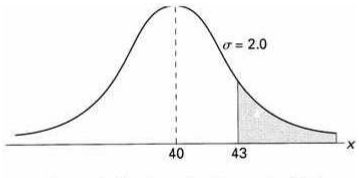

Figure 18: Area for Example 6.11.

- What is the lowest possible A and the highest possible B?
- Solution:

$$1 - 0.12 = 0.88 = P(Z < 1.175)$$
$$1.175 = \frac{x - 74}{7} \Rightarrow$$
$$x = 7 * 1.175 + 74 = 82.225$$

The lowest A is 83 and the highest B is 82.

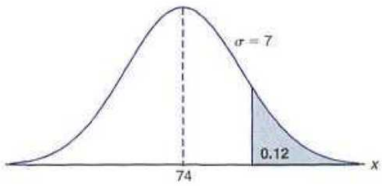

Figure 19: Area for Example 6.13.