Chapter 4, Section 5
Continuous Random Variables

Normal Distributions

Let $I = \int_{-\infty}^{\infty} e^{-x^2/2}dx$. (A multiple of this is the normal pdf.)

It is not hard to show that this integral exists, since $g(x) = e^{-x^2/2}$ is a positive, continuous function which is bounded above by an integrable function:

$$0 < e^{-x^2/2} < e^{-|x|+1} \quad \text{for} \quad -\infty < x < \infty$$

and $$\int_{-\infty}^{\infty} e^{-|x|+1}dx = 2\int_{0}^{\infty} e^{-x+1}dx = -2e^{-x+1}\Big|_0^{\infty} = 2e.$$

Since $\int_{-\infty}^{\infty} e^{-x^2/2}dx$ exists, so does $\int_{a}^{b} e^{-x^2/2}dx$ for all $a, b \in \mathbf{R}$.

However, it cannot be evaluated using elementary functions; we must use tables or numerical approximation methods.

To find the "standard" normal probability density function, a multiple of $g(x) = e^{-x^2/2}$, we need to evaluate $I = \int_{-\infty}^{\infty} e^{-x^2/2} dx$.

It involves a trick: evaluate I^2 by transforming to polar coordinates.

$$
\begin{aligned}
I^2 &= I \cdot I \\
&= \int_{-\infty}^{\infty} e^{-x^2/2} dx \int_{-\infty}^{\infty} e^{-y^2/2} dy \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-\left(x^2+y^2\right)/2} dy\, dx \\
&= \int_{0}^{2\pi} \int_{0}^{\infty} e^{-r^2/2}\, r\, dr\, d\theta
\end{aligned}
$$

Transform to polar coordinates:

$x = r \cos \theta$
$y = r \sin \theta$
$x^2 + y^2 = r^2$
$dA = dy\, dx = r\, dr\, d\theta$

and

$-\infty < x < \infty$, $-\infty < y < \infty$

is equivalent to

$0 \le r < \infty$, $0 \le \theta < 2\pi$

$$
I^2 = \int_{0}^{2\pi} \int_{0}^{\infty} e^{-r^2/2}\, r\, dr\, d\theta
$$

Substitute $u = -\frac{r^2}{2}$ for r:

$u = -\frac{r^2}{2}$
$du = -r\, dr$

r	$u = -\frac{r^2}{2}$
0	0
∞	$-\infty$

$$
\begin{aligned}
&= \int_0^{2\pi} \int_0^{\infty} \left(-e^{-r^2/2}\right)(-r)\, dr\, d\theta \\
&= \int_0^{2\pi} \int_0^{-\infty} \left(-e^{u}\right) du\, d\theta \\
&= \int_0^{2\pi} \int_{-\infty}^{0} e^{u}\, du\, d\theta \\
&= \int_0^{2\pi} e^{u}\Big|_{-\infty}^{0} d\theta \\
&= \int_0^{2\pi} (1-0)\, d\theta \\
&= \int_0^{2\pi} 1\, d\theta = \theta\Big|_0^{2\pi} = 2\pi - 0 = 2\pi.
\end{aligned}
$$

Thus, $I = \int_{-\infty}^{\infty} e^{-z^2/2} dz = \sqrt{2\pi}$, so $\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-z^2/2}\, dz = 1$ and

$f(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2}, \quad -\infty < z < \infty,$ is a density function, the

***N*(0, 1)-density function** or the **standard normal density function.**

Now let $z = \frac{y-\mu}{\sigma}$, where $\sigma > 0$ and μ is any real number. Then

$$1 = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-z^2/2}\, dz$$

$$= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(y-\mu)^2}{2\sigma^2}\right) dy\,.$$

$dz = \frac{1}{\sigma}\, dy$

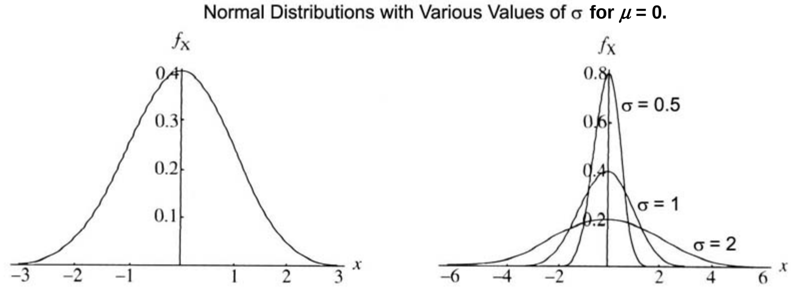
$y = \mu + \sigma z$

z	$y = \mu + \sigma z$
$-\infty$	$-\infty$
∞	∞

Thus $f(y) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(y-\mu)^2}{2\sigma^2}\right), \quad -\infty < y < \infty,$ is a density function,

that of the **$N(\mu, \sigma^2)$ – distribution** or the **Normal Distribution with parameters μ and σ^2**. Its graph is the "bell-shaped curve."

Normal Distributions with Various Values of σ **for μ = 0.**

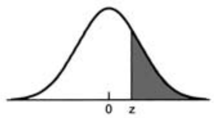


a. Normal (0, 1)

b. Normal (0, σ^2)

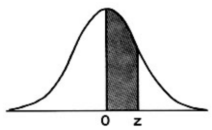
Note the symmetry about the vertical line through μ = 0.

Different Forms of the Standard Normal Distribution Table

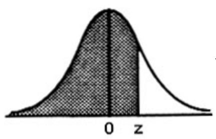


Our textbook

z	.00	.01	.02	.03	.04	.05
0.0	.5000	.4960	.4920	.4880	.4840	.4801
0.1	.4602	.4562	.4522	.4483	.4443	.4404
0.2	.4207	.4168	.4129	.4090	.4052	.4013
0.3	.3821	.3783	.3745	.3707	.3669	.3632
0.4	.3446	.3409	.3372	.3336	.3300	.3264
0.5	.3085	.3050	.3015	.2981	.2946	.2912
0.6	.2743	.2709	.2676	.2643	.2611	.2578
0.7	.2420	.2389	.2358	.2327	.2296	.2266
0.8	.2119	.2090	.2061	.2033	.2005	.1977
0.9	.1841	.1814	.1788	.1762	.1736	.1711

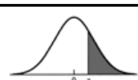


Some other texts



SOA Exam and TI-83

z	0.00	0.01	0.02	0.03	0.04	0.05
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289



z	.00	.01	.02	.03	.04	.05
0.0	.5000	.4960	.4920	.4880	.4840	.4801
0.1	.4602	.4562	.4522	.4483	.4443	.4404
0.2	.4207	.4168	.4129	.4090	.4052	.4013
0.3	.3821	.3783	.3745	.3707	.3669	.3632
0.4	.3446	.3409	.3372	.3336	.3300	.3264
0.5	.3085	.3050	.3015	.2981	.2946	.2912
0.6	.2743	.2709	.2676	.2643	.2611	.2578
0.7	.2420	.2389	.2358	.2327	.2296	.2266
0.8	.2119	.2090	.2061	.2033	.2005	.1977
0.9	.1841	.1814	.1788	.1762	.1736	.1711
1.0	.1587	.1562	.1539	.1515	.1492	.1469
1.1	.1357	.1335	.1314	.1292	.1271	.1251
1.2	.1151	.1131	.1112	.1093	.1075	.1056
1.3	.0968	.0951	.0934	.0918	.0901	.0885
1.4	.0808	.0793	.0778	.0764	.0749	.0735
1.5	.0668	.0655	.0643	.0630	.0618	.0606
1.6	.0548	.0537	.0526	.0516	.0505	.0495
1.7	.0446	.0436	.0427	.0418	.0409	.0401
1.8	.0359	.0352	.0344	.0336	.0329	.0322
1.9	.0287	.0281	.0274	.0268	.0262	.0256

Examples using the table.

If $Z \sim N(0, 1)$, find:
(first draw the pictures)

$P(Z \geq 1.84)$

$P(Z \leq 1.84)$

$P(Z \leq -1.84)$

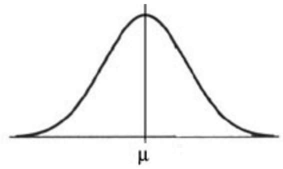
$P(-1.2 \leq Z \leq 1.84)$

$P(1.2 \leq Z \leq 1.84)$

c so that $P(Z \leq c) = 0.9357$

Properties of the Normal Density Functions. Let $Y \sim N(\mu, \sigma^2)$, so

$$f(y) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(y-\mu)^2}{2\sigma^2}\right), \quad -\infty < y < \infty.$$



1. The graph of f is symmetrical about the vertical line $y = \mu$, so $\mu = E(Y)$.

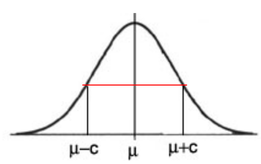
2. $f(\mu) = \frac{1}{\sqrt{2\pi}\sigma}$ is the absolute maximum of f on $(-\infty, \infty)$.

3. $f(y)$ has points of inflection at $y = \mu \pm \sigma$.

4. $\lim_{y \to -\infty} f(y) = 0 = \lim_{y \to +\infty} f(y)$.

Properties. Let $Y \sim N(\mu, \sigma^2)$, so $f(y) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(y-\mu)^2}{2\sigma^2}\right), \quad -\infty < y < \infty.$

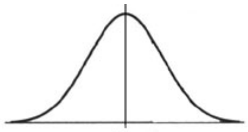
1. The graph of f is symmetrical about the line $y = \mu$.



Proof. y enters into the formula for $f(y)$ only through the expression $(y - \mu)^2$. Since $y - \mu$ is the signed distance from μ to y, the graph of $f(y)$ has the same height at the two points on either side of μ which are the same distance from μ.

Properties. Let $Y \sim N(\mu, \sigma^2)$, so $f(y) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(y-\mu)^2}{2\sigma^2}\right)$, $-\infty < y < \infty$.

2. $f(\mu) = \frac{1}{\sqrt{2\pi}\sigma}$ is the absolute maximum of f on $(-\infty, \infty)$.



Proof. y enters into the formula for $f(y)$ only through the expression $(y-\mu)^2$. Since the argument of the exponential, $\exp\left(-\frac{(y-\mu)^2}{2\sigma^2}\right)$, is negative unless $y = \mu$ and is zero when $y = \mu$, the density function, f, has its maximum value when $y = \mu$.

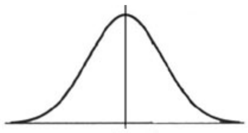
3. $f(y)$ has points of inflection at $y = \mu \pm \sigma$.

[This is an exercise in calculus.]

11

Properties. Let $Y \sim N(\mu, \sigma^2)$, so $f(y) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(y-\mu)^2}{2\sigma^2}\right)$, $-\infty < y < \infty$.

4. $\lim_{y\to-\infty} f(y) = 0 = \lim_{y\to+\infty} f(y)$.



Proof. y appears in the formula for $f(y)$ only in the expression

$\exp\left(-\frac{(y-\mu)^2}{2\sigma^2}\right)$ and $\lim_{y\to\pm\infty} \exp\left(-\frac{(y-\mu)^2}{2\sigma^2}\right) = 0$.

12

Theorem. If $Y \sim N(\mu, \sigma^2)$, then $E(Y) = \mu$ and $V(Y) = \sigma^2$.

(Since the density function of Y is symmetric about the vertical line $y = \mu$, we obtain $E(Y) = \mu$. The proof that $V(Y) = \sigma^2$ will be deferred until we discuss moment-generating functions.)

Theorem. If $Y \sim N(\mu, \sigma^2)$ and $U = aY + b$, where a and b are constants with $a \neq 0$, then U has a $N(a\mu + b, a^2\sigma^2)$–distribution.

Special Cases.

1. If $Z \sim N(0, 1)$ and $Y = \sigma Z + \mu$, then $Y \sim N(\mu, \sigma^2)$.
2. If $Y \sim N(\mu, \sigma^2)$ and $Z = \frac{Y-\mu}{\sigma} \left(= \frac{1}{\sigma}Y - \frac{\mu}{\sigma}\right)$, then $Z \sim N(0, 1)$.

[This is known as "**standardizing**" the normal random variable, Y.]

Proof of the theorem. Suppose that $Y \sim N(\mu, \sigma^2)$ and $U = aY + b$, with $a \neq 0$. We shall use the distribution function of Y to find the distribution function of U. This is a taste of part of Chapter 6.

Let $F(y)$ be the distribution function of Y. Then Y has density function

$$F'(y) = f(y) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(y-\mu)^2}{2\sigma^2}\right), -\infty < y < \infty.$$

Let $G(u)$ and $g(u)$ be the distribution and density functions, respectively, of U. Then, by the definition of a distribution function,

$$G(u) = P(U \leq u) = P(aY + b \leq u) = P(aY \leq u - b)$$

for all real numbers, u.

We need to consider two cases: $a > 0$ and $a < 0$.

Case 1. $a > 0$. Then $G(u) = \Pr(aY \le u-b) = \Pr\left(Y \le \frac{u-b}{a}\right) = F\left(\frac{u-b}{a}\right)$,

so
$$g(u) = G'(u) = F'\left(\frac{u-b}{a}\right) \cdot \frac{d}{du}\left(\frac{u-b}{a}\right)$$
$$= f\left(\frac{u-b}{a}\right) \cdot \left(\frac{1}{a}\right)$$
$$= \frac{1}{a} \cdot \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\left[\left(\frac{u-b}{a}\right) - \mu\right]^2 \Big/ \left(2\sigma^2\right)\right\}$$
$$= \frac{1}{\sqrt{2\pi}\, a\sigma} \exp\left\{-\left[\frac{u-b-a\mu}{a}\right]^2 \Big/ \left(2\sigma^2\right)\right\}$$
$$= \frac{1}{\sqrt{2\pi}\, a\sigma} \exp\left\{-\left[\frac{u-(a\mu+b)}{2a^2\sigma^2}\right]^2\right\},$$

which is the density of the $N(a\mu + b, a^2\sigma^2)$-distribution since

$\sqrt{a^2\sigma^2} = a\sigma$ when $a > 0$.

Case 2. $a < 0$. $G(u) = \Pr(aY \le u-b) = \Pr\left(Y \ge \frac{u-b}{a}\right) = 1 - F\left(\frac{u-b}{a}\right)$, so

$$g(u) = G'(u) = -F'\left(\frac{u-b}{a}\right) \cdot \frac{d}{du}\left(\frac{u-b}{a}\right)$$
$$= -f\left(\frac{u-b}{a}\right) \cdot \left(\frac{1}{a}\right)$$
$$= -\frac{1}{a} \cdot \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\left[\left(\frac{u-b}{a}\right) - \mu\right]^2 \Big/ \left(2\sigma^2\right)\right\}$$
$$= \frac{1}{\sqrt{2\pi}\, |a|\sigma} \exp\left\{-\left[\frac{u-b-a\mu}{a}\right]^2 \Big/ \left(2\sigma^2\right)\right\} \qquad \text{since } a < 0 \Rightarrow -a = |a|$$
$$= \frac{1}{\sqrt{2\pi}\, |a|\sigma} \exp\left\{-\left[\frac{u-(a\mu+b)}{2a^2\sigma^2}\right]^2\right\},$$

But since $a < 0$, $\sqrt{a^2\sigma^2} = |a|\sigma$, so in this case too,

U has a $N(a\mu + b, a^2\sigma^2)$-distribution.

Example (p. 183 #4.66, 4.67, and extensions). A machining operation produces bearings with diameters that are normally distributed with mean 3.0005″ and standard deviation 0.0010″. Specifications require the diameters to lie in the interval 3.0000 ± 0.0020 inches. Those outside this interval are considered scrap. With the current setting:

1. What fraction of the total production will be scrap?
2. What is the appropriate value of c so that the probability that the diameter of a randomly chosen bearing is less than c is 95%?
3. What is the probability that if four bearings are chosen at random, at least two will be scrap?
4. What should the mean diameter be in order that the fraction of bearings scrapped be minimized?

The diameter of a randomly chosen bearing, $Y \sim N(3.0005, (0.0010)^2)$
Scrap: diameter is outside interval 3.0000 ± 0.0020 = [2.9980, 3.0020]

1. What fraction of the total production will be scrap?

Want $P(Y > 3.0020 \text{ or } Y < 2.9980) = P(Y > 3.0020) + P(Y < 2.9980)$

since $(Y > 3.0020)$ and $(Y < 2.9980)$ are mutually exclusive events.

Standardize Y & use Normal Table: $Z = \frac{Y-\mu}{\sigma} = \frac{Y-3.0005}{0.0010} \sim N(0,1)$

$$P(Y > 3.0020) = P\left(\frac{Y-3.0005}{0.0010} > \frac{3.0020-3.0005}{0.0010}\right) = P(Z > 1.5) = 0.0668$$

$$P(Y < 2.9980) = P\left(\frac{Y-3.0005}{0.0010} < \frac{2.9980-3.0005}{0.0010}\right) = P(Z < -2.5)$$

$$= P(Z > 2.5) = 0.0062$$

Thus, $P(Y > 3.0020 \text{ or } Y < 2.9980) = 0.0668 + 0.0062 = 0.0730$.
In other words, about 7.3% of the output will be scrap

The diameter of a randomly chosen bearing, $Y \sim N(3.0005, (0.0010)^2)$

Scrap: diameter is outside interval $3.0000 \pm 0.0020 = [2.9980, 3.0020]$

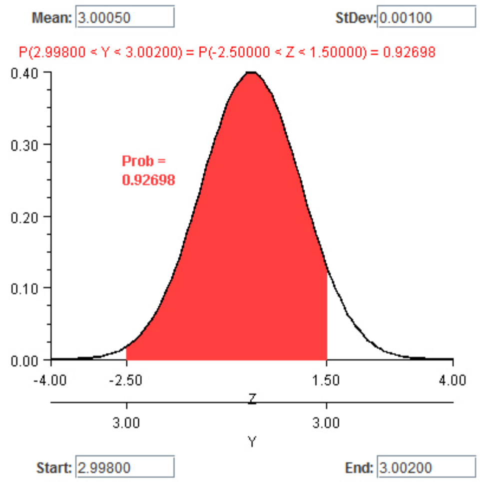
1. What fraction of the total production will be scrap?

We can also calculate $P(2.9980 \le Y \le 3.0020)$, the fraction that is not scrap, and subtract from 1.

Using the text's applet to do this, we get that

$1 - 0.92698 = 0.07302$

or about 7.302% of the total production is scrap.



The diameter of a randomly chosen bearing, $Y \sim N(3.0005, (0.0010)^2)$

2. What is the appropriate value of c so that the probability that the diameter of a randomly chosen bearing is less than c is 95%? That is, what is the 0.95 quantile or 95th percentile of the distribution of Y?

We need to find c so that $P(Y < c) = 0.95$.

Standardize Y & use Normal Table: $Z = \frac{Y - 3.0005}{0.0010} \sim N(0,1)$

So we need to solve

$$P\left(\frac{Y-3.0005}{0.0010} < \frac{c-3.0005}{0.0010}\right) = P\left(Z < \frac{c-3.0005}{0.0010}\right) = 0.95 \text{ for } c.$$

Because of the form of our tables, we need to think of this as solving

$$P\left(Z \ge \frac{c-3.0005}{0.0010}\right) = 1 - 0.95 = 0.05 \text{ for } c.$$

Solve $P(Z \geq \frac{c-3.0005}{0.0010}) = 0.05$ for c.

The value, z_0, for which $P(Z \geq z_0) = 0.05$ can be obtained by searching for 0.0500 in the body of the table.

0.0500 is not there, but we find

0.0495 (corresponding to $z = 1.65$), and

0.0505 (corresponding to $z = 1.64$).

Since 0.0500 is half-way between 0.0495 and 0.0505, we estimate z_0 by $\frac{1.65+1.64}{2} = 1.645$. Thus, $\frac{c-3.0005}{0.0010} = 1.645$

$\Rightarrow c - 3.0005 = 0.001645$

$\Rightarrow c = 3.002145 \approx 3.0021$ is the 95th percentile of Y.

The diameter of a randomly chosen bearing, $Y \sim N(3.0005, (0.0010)^2)$

3. What is the probability that if four bearings are chosen at random, at least two will be scrap?

Solution. Four bearings are selected at random. Let X = the number which are scrap. We are asked for $P(X \geq 2)$.

Since the four bearings are chosen at random from a large lot, X has an approximate binomial distribution with $n = 4$.

By part (a), the probability, p, of any one bearing's being scrap is 0.0730, so $X \approx \text{bin}(4, 0.073)$.

Notice that the random variable of interest here is X, the number among the 4 selected which are scrap, and X is binomial. We use the original random variable, Y, the diameter of a bearing, to determine the parameter, p, of the distribution of X.

The diameter of a randomly chosen bearing, $Y \sim N(3.0005, (0.0010)^2)$

3. What is the probability that if four bearings are chosen at random, at least two will be scrap?

Solution. Let X = the number of selected bearings which are scrap. We are asked for $P(X \geq 2)$ where $X \approx \text{bin}(4, 0.073)$.

$$P(X \geq 2) = 1 - P(X < 2) = 1 - P(X \leq 1) = 1 - P(X = 0) - P(X = 1)$$

$$= 1 - \binom{4}{0}(0.073)^0(0.927)^4 - \binom{4}{1}(0.073)^1(0.927)^3$$

$$= 1 - 0.738446 - 0.232607 = 0.0289.$$

Thus, the probability that at least two of the four chosen bearings are scrap is approximately 2.9%.

The diameter of a randomly chosen bearing, $Y \sim N(3.0005, (0.0010)^2)$

Scrap: diameter is outside interval $3.0000 \pm 0.0020 = [2.9980, 3.0020]$

4. What should the mean diameter be in order that the fraction of bearings scrapped be minimized?

Solution. In order to minimize the scrap fraction, we need the maximum amount in the specifications interval. Since the normal distribution is symmetric and unimodal, the mean diameter should be set to be the midpoint of the interval, or $\mu = 3.000$ in.

This is because for a symmetric, unimodal distribution, Y, the probability that Y is in an interval of a specified size (0.0040 here) is maximized when that interval is symmetric about the mean.