Chapter 3. Modeling Process Quality
Stem-and-Leaf Display

- Percentile
- Sample median or fiftieth percentile
- First quartile (Q1), third quartile (Q3)
- Interquartile range (Q3-Q1)
Plot of Data in Time Order

Figure 2-2  A time series plot of the health insurance data in Table 2-1.
# Histograms – Useful for large data sets

**Table 2-2**  Layer Thickness (Å) on Semiconductor Wafers

<p>| | | | | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>438</td>
<td>450</td>
<td>487</td>
<td>451</td>
<td>452</td>
<td>441</td>
<td>444</td>
<td>461</td>
<td>432</td>
<td>471</td>
</tr>
<tr>
<td>413</td>
<td>450</td>
<td>430</td>
<td>437</td>
<td>465</td>
<td>444</td>
<td>471</td>
<td>453</td>
<td>431</td>
<td>458</td>
</tr>
<tr>
<td>444</td>
<td>450</td>
<td>446</td>
<td>444</td>
<td>466</td>
<td>458</td>
<td>471</td>
<td>452</td>
<td>455</td>
<td>445</td>
</tr>
<tr>
<td>468</td>
<td>459</td>
<td>450</td>
<td>453</td>
<td>473</td>
<td>454</td>
<td>458</td>
<td>438</td>
<td>447</td>
<td>463</td>
</tr>
<tr>
<td>445</td>
<td>466</td>
<td>456</td>
<td>434</td>
<td>471</td>
<td>437</td>
<td>459</td>
<td>445</td>
<td>454</td>
<td>423</td>
</tr>
<tr>
<td>472</td>
<td>470</td>
<td>433</td>
<td>454</td>
<td>464</td>
<td>443</td>
<td>449</td>
<td>435</td>
<td>435</td>
<td>451</td>
</tr>
<tr>
<td>474</td>
<td>457</td>
<td>455</td>
<td>448</td>
<td>478</td>
<td>465</td>
<td>462</td>
<td>454</td>
<td>425</td>
<td>440</td>
</tr>
<tr>
<td>454</td>
<td>441</td>
<td>459</td>
<td>435</td>
<td>446</td>
<td>435</td>
<td>460</td>
<td>428</td>
<td>449</td>
<td>442</td>
</tr>
<tr>
<td>455</td>
<td>450</td>
<td>423</td>
<td>432</td>
<td>459</td>
<td>444</td>
<td>445</td>
<td>454</td>
<td>449</td>
<td>441</td>
</tr>
<tr>
<td>449</td>
<td>445</td>
<td>455</td>
<td>441</td>
<td>464</td>
<td>457</td>
<td>437</td>
<td>434</td>
<td>452</td>
<td>439</td>
</tr>
</tbody>
</table>

- Group values of the variable into bins, then count the number of observations that fall into each bin

- Plot frequency (or relative frequency) versus the values of the variable
Figure 2-3  Minitab histogram for the metal layer thickness data in Table 2-2.  

Figure 2-4  Minitab histogram with 15 bins for the metal layer thickness data.  

Figure 2-5  A cumulative frequency plot of the metal thickness data from Minitab.  

Figure 2-6  Histogram of the number of defects in painted automobile hoods (Table 2-3).
Table 2.3  Surface Finish Defects in Painted Automobile Hoods

| 6  | 1  | 5  | 7  | 8  | 6  | 0  | 2  | 4  | 2  |
| 5  | 2  | 4  | 4  | 1  | 4  | 1  | 7  | 2  | 3  |
| 4  | 3  | 3  | 3  | 6  | 3  | 2  | 3  | 4  | 5  |
| 5  | 2  | 3  | 4  | 4  | 4  | 2  | 3  | 5  | 7  |
| 5  | 4  | 5  | 5  | 4  | 5  | 3  | 3  | 3  | 12 |

Proportions of hoods with at least 3 defects = \( \frac{39}{50} = 0.78 \)

Proportions of hoods with between 0 and 2 defects = \( \frac{11}{50} = 0.22 \)
Numerical Summary of Data

Samples: $x_1, x_2, x_3, \ldots, x_n$

Sample average: \[ \bar{x} = \frac{x_1 + x_2 + \cdots + x_n}{n} = \frac{\sum_{i=1}^{n} x_i}{n} \]

Previous wafer thickness example:

\[ \bar{x} = \frac{\frac{\sum_{i=1}^{100} x_i}{100}}{100} = \frac{45.001}{100} = 450.01 \text{Å} \]
Sample Variance:

\[ s^2 = \frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n - 1} \]

Sample Standard Deviation:

\[ s = \sqrt{\frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n - 1}} \]

Previous wafer thickness example:

\[ s^2 = 180.2928 \text{ Å}^2 \]

\[ s = 13.43 \text{ Å} \]
The Box Plot
(or Box-and-Whisker Plot)

Table 2-4  Hole Diameters (in mm) in Wing Leading Edge Ribs

<table>
<thead>
<tr>
<th>Hole Diameter</th>
<th>120.5</th>
<th>120.9</th>
<th>120.3</th>
<th>121.3</th>
</tr>
</thead>
<tbody>
<tr>
<td>120.4</td>
<td>120.2</td>
<td>120.1</td>
<td>120.5</td>
<td></td>
</tr>
<tr>
<td>120.7</td>
<td>121.1</td>
<td>120.9</td>
<td>120.8</td>
<td>121.3</td>
</tr>
</tbody>
</table>

**Figure 2-7**  Box plot for the aircraft wing leading edge hole diameter data in Table 2-4.

Box plots can identify potential outliers.
Comparative Box Plots

Figure 2-8 Comparative box plots of a quality index for products produced at three plants.
Probability Distributions

- Statistics: based on sample analysis involving measurements
- Probability: based on mathematical description of abstract model
- Random variable: (real) value associated with variable of interest
- Probability distribution: probability of occurrence of variable

**Definition**

1. **Continuous distributions.** When the variable being measured is expressed on a continuous scale, its probability distribution is called a *continuous distribution*. The probability distribution of metal layer thickness is continuous.

2. **Discrete distributions.** When the parameter being measured can only take on certain values, such as the integers 0, 1, 2, . . . , the probability distribution is called a *discrete distribution*. For example, the distribution of the number of non-conformities or defects in printed circuit boards would be a discrete distribution.
Sometimes called a probability mass function

Sometimes called a probability density function

Figure 2-9  Probability distributions. (a) Discrete case. (b) Continuous case.
EXAMPLE 2-5  

A Discrete Distribution

A manufacturing process produces thousands of semiconductor chips per day. On the average, 1% of these chips do not conform to specifications. Every hour, an inspector selects a random sample of 25 chips and classifies each chip in the sample as conforming or nonconforming. If we let $x$ be the random variable representing the number of nonconforming chips in the sample, then the probability distribution of $x$ is

$\binom{25}{x} = \frac{25!}{x!(25-x)!}$

Error. Should be

$$p(x) = \binom{25}{x} (0.01)^x (0.99)^{25-x}$$

$x = 0, 1, 2, \ldots, 25$

where $\binom{25}{x} = 25!/[x!(25-x)!]$. This is a discrete distribution, since the observed number of nonconformances is $x = 0, 1, 2, \ldots, 25$, and is called the binomial distribution. We may calculate the probability of finding one or fewer nonconforming parts in the sample as

$$P(x \leq 1) = P(x = 0) + P(x = 1)$$

$$= p(0) + p(1)$$

$$= \sum_{x=0}^{1} \binom{25}{x} (0.01)^x (0.99)^{25-x}$$

$$= \frac{25!}{0!25!} (0.01)^0 (0.99)^{25} + \frac{25!}{1!24!} (0.01)^1 (0.99)^{24}$$

$$= 0.7778 + 0.1964 = 0.9742$$
EXAMPLE 2-6

A Continuous Distribution

Suppose that \( x \) is a random variable that represents the actual contents in ounces of a 1-lb bag of coffee beans. The probability distribution of \( x \) is assumed to be

\[
f(x) = \frac{1}{1.5} \quad 15.5 \leq x \leq 17.0
\]

This is a continuous distribution, since the range of \( x \) is the interval \([15.5, 17.0]\). This distribution is called the uniform distribution, and it is shown graphically in Fig. 2-10. Note that the area under the function \( f(x) \) corresponds to probability, so that the probability of a bag containing less than 16.0 oz is

\[
P\{x \leq 16.0\} = \int_{15.5}^{16.0} f(x) \, dx = \int_{15.5}^{16.0} \frac{1}{1.5} \, dx = \left. \frac{x}{1.5} \right|_{15.5}^{16.0} = \frac{16.0 - 15.5}{1.5} = 0.3333
\]

Figure 2-10 The uniform distribution for Example 2-6.
Mean
\[ \mu = \begin{cases} 
\int_{-\infty}^{\infty} xf(x) \, dx, & x \text{ continuous} \\
\sum_{i=1}^{\infty} x_i p(x_i), & x \text{ discrete}
\end{cases} \]

Variance
\[ \sigma^2 = \begin{cases} 
\int_{-\infty}^{\infty} (x - \mu)^2 f(x) \, dx, & x \text{ continuous} \\
\sum_{i=1}^{\infty} (x_i - \mu)^2 p(x_i), & x \text{ discrete}
\end{cases} \]
Figure 2-11  The mean of a distribution.

Figure 2-12  Two probability distributions with different means.

Figure 2-13  Two probability distributions with the same mean but different standard deviations.
\( N: \) population
\( D: \) class of interest within population
\( n: \) random samples chosen from \( N \)
\( x: \) belongs to class of interest among random samples

**Definition**

The **hypergeometric probability distribution** is

\[
p(x) = \binom{D}{x} \binom{N-D}{n-x} \binom{N}{n} \quad x = 0, 1, 2, \ldots, \min(n, D) \quad (2-8)
\]

The mean and variance of the distribution are

\[
\mu = \frac{nD}{N} \quad (2-9)
\]

and

\[
\sigma^2 = \frac{nD}{N} \left(1 - \frac{D}{N}\right) \left(\frac{N-n}{N-1}\right) \quad (2-10)
\]
Example: A lot containing 100 products. There are five defects. Choose 10 random samples. Find the probability that one or fewer defects are contained in the sample.

\[ N = 100 \]
\[ D = 5 \]
\[ n = 10 \]
\[ x \leq 1 \]

\[
P\{x \leq 1\} = P\{x = 0\} + P\{x = 1\}
\]
\[
= \frac{\binom{5}{0}\binom{95}{10}}{\binom{100}{10}} + \frac{\binom{5}{1}\binom{95}{9}}{\binom{100}{10}} = 0.923
\]
The random variable $x$ is the number of successes out of $n$ independent Bernoulli trials with constant probability of success $p$ on each trial.

**Definition**

The **binomial distribution** with parameters $n \geq 0$ and $0 < p < 1$ is

$$p(x) = \binom{n}{x} p^x (1 - p)^{n-x} \quad x = 0, 1, \ldots, n$$  \hspace{1cm} (2-11)

The mean and variance of the binomial distribution are

$$\mu = np$$  \hspace{1cm} (2-12)

and

$$\sigma^2 = np(1 - p)$$  \hspace{1cm} (2-13)
Example: $n = 15$, $p = 0.1$:

<table>
<thead>
<tr>
<th>$x$</th>
<th>$p(x)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.2059</td>
</tr>
<tr>
<td>1</td>
<td>0.3432</td>
</tr>
<tr>
<td>2</td>
<td>0.2669</td>
</tr>
<tr>
<td>3</td>
<td>0.1285</td>
</tr>
<tr>
<td>4</td>
<td>0.0428</td>
</tr>
<tr>
<td>5</td>
<td>0.0105</td>
</tr>
<tr>
<td>6</td>
<td>0.0019</td>
</tr>
<tr>
<td>7</td>
<td>0.0003</td>
</tr>
<tr>
<td>8</td>
<td>0.0001</td>
</tr>
<tr>
<td>15</td>
<td>0.0000</td>
</tr>
</tbody>
</table>
Error: This should be $p(x)$.

Figure 2-14  Binomial distributions for selected values of $n$ and $p$. 
Sample fraction defective

\[ \hat{p} = \frac{x}{n} \]

\( \hat{p} \) is an estimator of \( p \).

\[
P\{\hat{p} \leq a\} = P\left\{ \frac{x}{n} \leq a \right\} = p\{x \leq na\} = \sum_{x=0}^{[na]} \binom{n}{x} p^x (1 - p)^{n-x}
\]

\([na]\) is the largest integer less than or equal to \( na \).

Mean of \( \hat{p} = p \). Variance of \( \hat{p} = \sigma_{\hat{p}}^2 = \frac{p(1-p)}{n} \)
The **Poisson distribution** is

\[ p(x) = \frac{e^{-\lambda} \lambda^x}{x!} \quad x = 0, 1, \ldots \]  

(2-15)

where the parameter \( \lambda > 0 \). The **mean** and **variance** of the Poisson distribution are

\[ \mu = \lambda \]  

(2-16)

and

\[ \sigma^2 = \lambda \]  

(2-17)
As $\lambda$ gets large, $p(x)$ looks more symmetric, i.e., looks like binomial. Binomial is an approximation for limiting Poisson distribution. $n \to \infty$ and $p \to 0$ such that $np = \lambda$, binomial distribution approximates Poisson distribution with $\lambda$. 

**Figure 2-15** Poisson probability distributions for selected values of $\lambda$. 
Definition

The **Pascal distribution** is

\[ p(x) = \binom{x-1}{r-1} p^r (1-p)^{x-r} \quad x = r, r+1, r+2, \ldots \quad (2-18) \]

where \( r \geq 1 \) is an integer. The **mean** and **variance** of the Pascal distribution are

\[ \mu = \frac{r}{p} \quad (2-19) \]

and

\[ \sigma^2 = \frac{r(1-p)}{p^2} \quad (2-20) \]

respectively.

The random variable \( x \) is the number of Bernoulli trials upon which the \( r \)th success occurs.
• When $r = 1$ the Pascal distribution is known as the geometric distribution.

• The geometric distribution has many useful applications in statistical quality control.
The normal distribution is

\[ f(x) = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{1}{2} \left( \frac{x-\mu}{\sigma} \right)^2}, \quad -\infty < x < \infty \]  

(2-21)

The mean of the normal distribution is \( \mu \) (\( -\infty < \mu < \infty \)) and the variance is \( \sigma^2 > 0 \).

\[ x \sim N(\mu, \sigma^2) \]: \( x \) is normally distributed with mean \( \mu \) and variance \( \sigma^2 \).
where

\[ Z = \frac{X - \mu}{\sigma} \]

and \( \Phi(.) \) is a standard normal distribution with mean 0 and standard deviation 1.
EXAMPLE 2-7

The tensile strength of paper used to make grocery bags is an important quality characteristic. It is known that the strength—say, \( x \)—is normally distributed with mean \( \mu = 40 \text{ lb/in}^2 \) and standard deviation \( \sigma = 2 \text{ lb/in}^2 \), denoted \( x \sim N(40, 2^2) \). The purchaser of the bags requires them to have a strength of at least 35 lb/in\(^2\). The probability that a bag produced from this paper will meet or exceed this specification is \( P\{x \geq 35\} \). Note that

\[
P\{x \geq 35\} = 1 - P\{x \leq 35\}
\]

To evaluate this probability from the standard normal tables, we standardize the point 35 and find

\[
P\{x \leq 35\} = P\left\{ z \leq \frac{35 - 40}{2} \right\} = P\{ z \leq -2.5 \} = \Phi(-2.5) = 0.0062
\]

Consequently, the desired probability is

\[
P\{x \geq 35\} = 1 - P\{x \leq 35\} = 1 - 0.0062 = 0.9938
\]

Figure 2-18 shows the tabulated probability for both the \( N(40, 2^2) \) distribution and the standard normal distribution. Note that the shaded area to the left of 35 lb/in\(^2\) in Fig. 2-18 represents the fraction nonconforming or “fallout” produced by the bag manufacturing process.
Figure 2-18  Calculation of $P\{x \leq 35\}$ in Example 2-7.
EXAMPLE 2.9

Sometimes instead of finding the probability associated with a particular value of a normal random variable, we find it necessary to do the opposite—find a particular value of a normal random variable that results in a given probability. For example, suppose that \( x \sim N(10, 9) \). We wish to find the value of \( x \)—say, \( a \)—such that \( P \{ x > a \} = 0.05 \). Thus

\[
P \{ x > a \} = P \left\{ z > \frac{a - 10}{3} \right\} = 0.05
\]

or

\[
P \left\{ z \leq \frac{a - 10}{3} \right\} = 0.95
\]

From Appendix Table II, we have \( P \{ z \leq 1.645 \} = 0.95 \), so

\[
\frac{a - 10}{3} = 1.645
\]

or

\[
a = 10 + 3(1.645) = 14.935
\]
Let $x_i \sim N\left(\mu_i, \sigma_i^2\right)$ and $x_i$ for $i = 1, 2, \ldots, n$ are independent.

Then, $y = a_1 x_1 + a_2 x_2 + a_3 x_3 + \ldots + a_n x_n$

$\sim N\left(a_1 \mu_1 + a_2 \mu_2 + a_3 \mu_3 + \ldots + a_n \mu_n, a_1^2 \sigma_1^2 + a_2^2 \sigma_2^2 + a_3^2 \sigma_3^2 + \ldots + a_n^2 \sigma_n^2\right)$
Definition: The Central Limit Theorem

If \( x_1, x_2, \ldots, x_n \) are independent random variables with mean \( \mu_i \) and variance \( \sigma_i^2 \), and if \( y = x_1 + x_2 + \ldots + x_n \), then the distribution of

\[
\frac{y - \sum_{i=1}^{n} \mu_i}{\sqrt{\sum_{i=1}^{n} \sigma_i^2}}
\]

approaches the \( N(0, 1) \) distribution as \( n \) approaches infinity.

- Practical interpretation – the sum of independent random variables is approximately normally distributed regardless of the distribution of each individual random variable in the sum
Definition

Let \( w \) have a normal distribution mean \( \theta \) and variance \( \omega^2 \); then \( x = \exp(w) \) is a lognormal random variable, and the lognormal distribution is

\[
f(x) = \frac{1}{x\omega\sqrt{2\pi}} \exp\left[-\frac{(\ln(x) - \theta)^2}{2\omega^2}\right] \quad 0 < x < \infty \quad (2-29)
\]

The mean and variance of \( x \) are

\[
\mu = e^{\theta + \omega^2/2} \quad \text{and} \quad \sigma^2 = e^{2\theta + \omega^2} \left(e^{\omega^2} - 1\right) \quad (2-30)
\]
Figure 2-20  Lognormal probability density functions with $\theta = 0$ for selected values of $\omega^2$. 
EXAMPLE 2-10

The lifetime of a medical laser used in ophthalmic surgery has a lognormal distribution with $\theta = 6$ and $\omega = 1.2$ hours. What is the probability the lifetime exceeds 500 hours?

From the cumulative distribution function for the lognormal random variable

$$P(x > 500) = 1 - P[\exp(w) \leq 500] = 1 - P[w \leq \ln(500)]$$

$$= \Phi\left( \frac{\ln(500) - 6}{1.2} \right) = 1 - \Phi(0.1788) = 1 - 0.5710 = 0.4290$$

What lifetime is exceeded by 99% of lasers? Now the question is to determine $a$ such that $P(x > a) = 0.99$. Therefore,

$$P(x > a) = P[\exp(w) > a] = P[w > \ln(a)] = 1 - \Phi\left( \frac{\ln(a) - 6}{1.2} \right) = 0.99$$

From Appendix Table II, $1 - \Phi(a) = 0.99$ when $a = -2.33$. Therefore,

$$\frac{\ln(a) - 6}{1.2} = -2.33 \quad \text{and} \quad a = \exp(3.204) = 24.63 \text{ hours}$$

Determine the mean and standard deviation of lifetime. Now,

$$\mu = e^{\theta + \omega^2/2} = \exp(6 + 0.72) = 828.82 \text{ hours}$$

$$\sigma^2 = e^{2\theta + \omega^4/2} \{e^{\omega^2} - 1\} = \exp(12 + 1.44) \{\exp(1.44) - 1\} = 2,212,419.85$$

so the standard deviation of the lifetime is 1487.42 hours. Notice that the standard deviation of the lifetime is large relative to the mean.
Definition

The exponential distribution is

\[ f(x) = \lambda e^{-\lambda x} \quad x \geq 0 \]  

(2-31)

where \( \lambda > 0 \) is a constant. The mean and variance of the exponential distribution are

\[ \mu = \frac{1}{\lambda} \]  

(2-32)

and

\[ \sigma^2 = \frac{1}{\lambda^2} \]  

(2-33)

respectively.
**Figure 2-21** Exponential distributions for selected values of $\lambda$.

**Figure 2-22** The cumulative exponential distribution function.
The **gamma distribution** is

\[ f(x) = \frac{\lambda}{\Gamma(r)} (\lambda x)^{r-1} e^{-\lambda x} \quad x \geq 0 \quad (2-36) \]

with **shape parameter** \( r > 0 \) and **scale parameter** \( \lambda > 0 \). The **mean** and **variance** of the gamma distribution are

\[ \mu = \frac{r}{\lambda} \quad (2-37) \]

and

\[ \sigma^2 = \frac{r}{\lambda^2} \quad (2-38) \]

respectively.\(^3\)

---

\(^3\)\(\Gamma(r)\) in the denominator of equation 2-36 is the gamma function, defined as \( \Gamma(r) = \int_0^\infty x^{r-1} e^{-x} \, dx, \, r > 0 \). If \( r \) is a positive integer, then \( \Gamma(r) = (r - 1)! \).
• When \( r \) is an integer, the gamma distribution is the result of summing \( r \) independently and identically exponential random variables each with parameter \( \lambda \)

![Graph showing gamma distributions for selected values of \( r \) and \( \lambda = 1 \).](image)

**Figure 2-23** Gamma distributions for selected values of \( r \) and \( \lambda = 1 \).

![Diagram of a standby redundant system.](image)

**Figure 2-24** The standby redundant system for Example 2-11.
The **Weibull distribution** is

\[ f(x) = \frac{\beta}{\theta} \left( \frac{x}{\theta} \right)^{\beta-1} \exp \left[ -\left( \frac{x}{\theta} \right)^{\beta} \right] \quad x \geq 0 \quad (2-41) \]

where \( \theta > 0 \) is the **scale parameter**, and \( \beta > 0 \) is the **shape parameter**. The **mean** and **variance** of the Weibull distribution are

\[ \mu = \theta \Gamma \left( 1 + \frac{1}{\beta} \right) \quad (2-42) \]

and

\[ \sigma^2 = \theta^2 \left[ \Gamma \left( 1 + \frac{2}{\beta} \right) - \left\{ \Gamma \left( 1 + \frac{1}{\beta} \right) \right\}^2 \right] \quad (2-43) \]

respectively.
• When $\beta = 1$, the Weibull distribution reduces to the exponential distribution.

Figure 2-25  Weibull distributions for selected values of the shape parameter $\beta$ and scale parameter $\theta = 1$. 
Probability Plot

• Determining if a sample of data might reasonably be assumed to come from a specific distribution
• Probability plots are available for various distributions
• Easy to construct with computer software (MINITAB)
• Subjective interpretation
Observations on the road octane number of ten gasoline blends are as follows: 88.9, 87.0, 90.0, 88.2, 87.2, 87.4, 87.8, 89.7, 86.0, and 89.6. We hypothesize that octane number is adequately modeled by a normal distribution. To use probability plotting to investigate this hypothesis, first arrange the observations in ascending order and calculate their cumulative frequencies \((j - 0.5)/10\) as shown in the accompanying table.

<table>
<thead>
<tr>
<th>(j)</th>
<th>(x_{(j)})</th>
<th>((j - 0.5)/10)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>86.0</td>
<td>0.05</td>
</tr>
<tr>
<td>2</td>
<td>87.0</td>
<td>0.15</td>
</tr>
<tr>
<td>3</td>
<td>87.2</td>
<td>0.25</td>
</tr>
<tr>
<td>4</td>
<td>87.4</td>
<td>0.35</td>
</tr>
<tr>
<td>5</td>
<td>87.8</td>
<td>0.45</td>
</tr>
<tr>
<td>6</td>
<td>88.2</td>
<td>0.55</td>
</tr>
<tr>
<td>7</td>
<td>88.9</td>
<td>0.65</td>
</tr>
<tr>
<td>8</td>
<td>89.6</td>
<td>0.75</td>
</tr>
<tr>
<td>9</td>
<td>89.7</td>
<td>0.85</td>
</tr>
<tr>
<td>10</td>
<td>90.0</td>
<td>0.95</td>
</tr>
</tbody>
</table>

**Figure 2-26** Normal probability plot of the road octane number data.
Figure 2-27  Normal probability plot of the road octane number data with standardized scores.
Other Probability Plots

• What is a reasonable choice as a probability model for these data?

Table 2-5  Aluminum Contamination (ppm)

<p>| | | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>30</td>
<td>30</td>
<td>60</td>
<td>63</td>
<td>70</td>
<td>79</td>
<td>87</td>
<td></td>
</tr>
<tr>
<td>90</td>
<td>101</td>
<td>102</td>
<td>115</td>
<td>118</td>
<td>119</td>
<td>119</td>
<td></td>
</tr>
<tr>
<td>120</td>
<td>125</td>
<td>140</td>
<td>145</td>
<td>172</td>
<td>182</td>
<td></td>
<td></td>
</tr>
<tr>
<td>183</td>
<td>191</td>
<td>222</td>
<td>244</td>
<td>291</td>
<td>511</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 2-28  Probability plots of the aluminum contamination data in Table 2-5. (a) Normal. (b) Lognormal. (c) Weibull. (d) Exponential.
Approximations:  H = hypergeometric, B = binomial, P = Poisson, N = normal

Figure 2-29  Approximations to probability distributions