Simple Linear Regression I

Linear Regression

Linear regression: modeling the dependency of two variables using a linear approach.

The term was originally used in the sense of ***regression to the mean*** in biology. It was observed that certain extreme values of biological features in members of a population are not necessarily passed on to descendants, but that the descendants' values of these features return to being closer to the mean.

However, after continued analysis of the concrete biological problem, the term "regression" came to be used much more generally.

Setting and Assumptions

We have:

- a ***dependent variable Y***, which we will assume to be a random variable following a normal distribution. Y is often called the ***response variable***.
- an ***independent variable X***, which we can assume to either be a non-random parameter or a random variable measured precisely, without any error or uncertainty. X is often called the ***predictor variable*** or ***regressor***.

We want to describe the behavior of Y as a function of the values of X, i.e., of $Y \mid x$. We will therefore assume that there exists a certain ***model***.

For most f this discussion, we will take the point of view that x is not random while Y is a random variable following a normal distribution.

Simple Linear Regression Model

In this section, we assume that the mean $\mu_{Y|x}$ of $Y \mid x$ is given by

$$\mu_{Y|x} = \beta_0 + \beta_1 x \qquad \text{for some } \beta_0, \beta_1 \in \mathbb{R}. \tag{24.1}$$

This is called a ***simple linear regression*** model with ***model parameters*** β_0 and β_1.

Another way of writing this model is

$$Y \mid x = \beta_0 + \beta_1 x + E$$

where $\mathrm{E}[E] = 0$.

Our goal is to find estimators

$$B_0 := \widehat{\beta}_0 = \text{estimator for } \beta_0, \qquad b_0 = \text{estimate for } \beta_0,$$
$$B_1 := \widehat{\beta}_1 = \text{estimator for } \beta_1, \qquad b_1 = \text{estimate for } \beta_1,$$

Residuals

We assume that we have a random sample of size n of pairs (X, Y) or (if we consider X to be a parameter and not a random variable) $(x, Y \mid x)$.

For short, we write

$$Y_i := Y \mid x_i, \qquad\qquad i = 1, \ldots, n,$$

so that we have a random sample $(x_1, Y_1), \ldots, (x_n, Y_n)$.

According to our model, for each measurement y_i there exists a number e_i, called the ***residual***, such that

$$Y_i = b_0 + b_1 x_i + e_i.$$

Our goal is to determine b_0 and b_1 based on minimizing the residuals in a certain way.

Least-Squares Estimation

In 1805, Adrien Legendre published an approach for minimizing the residuals by letting

$$e_1^2 + e_2^2 + \cdots + e_n^2 \longrightarrow \text{minimum}.$$

Gauß published the same method (with a deeper analysis) in 1809 but claimed he had been using it since 1795. This set off a bitter priority dispute between Legendre and Gauß.

Adrien-Marie Legendre (17521833). Boilly, Julien-Leopold. (1820). Album de 73 Portraits-Charge Aquarelles des Membres de IInstitute (watercolor portrait # 29). Bibliotheque de IInstitut de France.

In a letter, Gauß notes that previously Laplace had been using the approach

$$|e_1| + |e_2| + \cdots + |e_n| \longrightarrow \text{minimum}$$

under the condition that $e_1 + e_1 + \cdots + e_n = 0$.

Gauß then extensively analyzed the ***least-squares method***.

Least Squares Estimation

Given a sample of size n, we define the ***error sum of squares***

$$\mathrm{SS}_\mathrm{E} := e_1^2 + e_2^2 + \cdots + e_n^2 = \sum_{i=1}^{n} \big(y_i - (b_0 + b_1 x_i)\big)^2 .$$

Since we determine the estimators for β_0 and β_1 by minimizing this sum of squares, b_0 and b_1 are called ***least-squares estimates***.

Assuming that $Y \mid x$ follows a normal distribution with variance σ^2 (independent of x) and mean $b_0 + b_1 x$, Gauß proved that the least-squares estimators have the smallest possible variance among all unbiased estimators for b_0 and b_1.

The Normal Equations

We consider SS_E as a function of b_0 and b_1 and find the minimum by calculating the partial derivatives:

$$\frac{\partial\, \text{SS}_\text{E}}{\partial b_0} = -2\sum_{i=1}^{n}(y_i - b_0 - b_1 x_i),$$

$$\frac{\partial\, \text{SS}_\text{E}}{\partial b_1} = -2\sum_{i=1}^{n}(y_i - b_0 - b_1 x_i)x_i$$

Setting the derivatives equal to zero, we obtain the so-called ***normal equations***

$$nb_0 + b_1\sum_{i=1}^{n} x_i = \sum_{i=1}^{n} y_i, \qquad b_0\sum_{i=1}^{n} x_i + b_1\sum_{i=1}^{n} x_i^2 = \sum_{i=1}^{n} x_i y_i$$

The Least Squares Estimates

These are linear equations for b_0 and b_1, which may be easily solved:

$$b_1 = \frac{n \sum_{i=1}^{n} x_i y_i - \left(\sum_{i=1}^{n} x_i\right)\left(\sum_{i=1}^{n} y_i\right)}{n \sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2}, \tag{24.2a}$$

$$b_0 = \frac{1}{n} \sum_{i=1}^{n} y_i - b_1 \cdot \frac{1}{n} \sum_{i=1}^{n} x_i. \tag{24.2b}$$

Although these formulas are straightforward for explicit calculations, it is worth re-writing them a little.

We will use the usual notation

$$\overline{x} = \frac{1}{n} \sum_{i=1}^{n} x_i \qquad \text{and} \qquad \overline{y} = \frac{1}{n} \sum_{i=1}^{n} y_i.$$

The Least Squares Estimates

Then it is easy to see that

$$\begin{aligned}\sum_{i=1}^{n}(x_i-\overline{x})(y_i-\overline{y}) &= \sum_{i=1}^{n} x_i y_i - \overline{y}\sum_{i=1}^{n} x_i - \overline{x}\sum_{i=1}^{n} y_i + n\cdot\overline{x}\cdot\overline{y} \\ &= \frac{1}{n}\Big(n\sum_{i=1}^{n} x_i y_i - \Big(\sum_{i=1}^{n} x_i\Big)\Big(\sum_{i=1}^{n} y_i\Big)\Big)\end{aligned}$$

For short, we will write

$$S_{xx} := \sum_{i=1}^{n}(x_i-\overline{x})^2, \qquad S_{yy} := \sum_{i=1}^{n}(y_i-\overline{y})^2,$$

$$S_{xy} := \sum_{i=1}^{n}(x_i-\overline{x})(y_i-\overline{y}).$$

The Least Squares Estimates

Then we can write

$$b_0 = \overline{y} - b_1 \overline{x}, \qquad\qquad b_1 = \frac{S_{xy}}{S_{xx}}.$$

24.1. Example. Since humidity influences evaporation, the solvent balance of water-reducible paints during sprayout is affected by humidity. A controlled study is conducted to examine the relationship between humidity (X) and the extent of solvent evaporation (Y). Knowledge of this relationship will be useful in that it will allow the painter to adjust his or her spray gun setting to account for humidity. The following data are obtained:

Linear Regression

Observation	x	y	Observation	x	y
1	35.3	11.0	14	39.1	9.6
2	27.7	11.1	15	46.8	10.9
3	30.8	12.5	16	48.5	9.6
4	58.8	8.4	17	59.3	10.1
5	61.4	9.3	18	70.0	8.1
6	71.3	8.7	19	70.0	6.8
7	74.4	6.4	20	74.4	8.9
8	76.7	8.5	21	72.1	7.7
9	70.7	7.8	22	58.1	8.5
10	57.5	9.1	23	44.6	8.9
11	46.4	8.2	24	33.4	10.4
12	28.9	12.2	25	28.6	11.1
13	28.1	11.9			

x is the observed relative humidity (in %), y is the observed solvent evaporation (in %).

Linear Regression

We obtain

$$n = 25, \qquad \sum_{i=1}^{n} x_i = 1312.9, \qquad \sum_{i=1}^{n} y_i = 235.70,$$

$$\sum_{i=1}^{n} x_i^2 = 76193.7, \qquad \sum_{i=1}^{n} y_i^2 = 2286.07, \qquad \sum_{i=1}^{n} x_i y_i = 11802.2$$

Then, using the formulas (24.2), we have

$$b_1 = \widehat{\beta}_1 = -0.0795, \qquad b_0 = \widehat{\beta}_0 = 13.6.$$

Hence the estimated regression equation is

$$\widehat{\mu}_{Y|x} = 13.6 - 0.0795x.$$

For example, the mean solvent evaporation at 50% relative humidity is estimated to be 9.63%.

Linear Regression with Mathematica

We use the data from Example 24.1 to illustrate how linear regression is implemented:

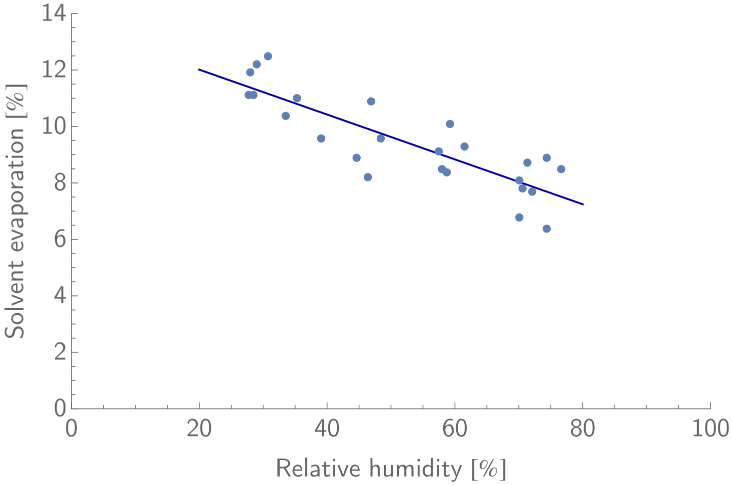
```
data := {{35.3, 11.0}, {27.7, 11.1}, {30.8, 12.5}, {58.8, 8.4},
   {61.4, 9.3}, {71.3, 8.7}, {74.4, 6.4}, {76.7, 8.5},
   {70.7, 7.8}, {57.5, 9.1}, {46.4, 8.2}, {28.9, 12.2},
   {28.1, 11.9}, {39.1, 9.6}, {46.8, 10.9}, {48.5, 9.6},
   {59.3, 10.1}, {70.0, 8.1}, {70.0, 6.8}, {74.4, 8.9},
   {72.1, 7.7}, {58.1, 8.5}, {44.6, 8.9}, {33.4, 10.4},
   {28.6, 11.1}};
model = LinearModelFit[data, x, x]

FittedModel[ 13.6013 - 0.0794677 x ]
```

Data is entered as a list of pairs (x_i, y_i) and the **LinearModelFit** command takes as its arguments the data, the model (here: a linear model in x) and the name of the variable (x).

Linear Regression with Mathematica

```
model["BestFit"]
```

```
13.6013 - 0.0794677 x
```



Model Assumptions and Random Samples

24.2. Model Assumption.

(i) For each value of x, the random variable $Y \mid x$ follows a normal distribution with variance σ^2 and mean $\mu_{Y|x} = \beta_0 + \beta_1 x$.

(ii) The random variables $Y \mid x_1$ and $Y \mid x_2$ are independent if $x_1 \neq x_2$.

A random sample of size n consists of n pairs (x_i, Y_i), $i = 1, \ldots, n$, where the random variables $Y_i = Y \mid x_i$ are i.i.d normal with variance σ^2 and mean $\mu_{Y|x_i} = \beta_0 + \beta_1 x_i$.

24.3. Remark. We do not require that $x_i \neq x_j$. The random sample may contain more than a single measurement of $Y \mid x_i$. All the x_i are treated in the same way, e.g., when calculating $\overline{x}$.

Distribution of the Least Squares Estimators

24.4. Theorem. Given a random sample of $Y \mid x$ of size n, the statistics

$$\frac{B_1 - \beta_1}{\sigma / \sqrt{\sum (x_i - \overline{x})^2}} \qquad \text{and} \qquad \frac{B_0 - \beta_0}{\sigma \sqrt{\frac{\sum x_i^2}{n \sum (x_i - \overline{x})^2}}}$$

follow a standard normal distribution.

In particular, B_0 and B_1 are unbiased estimators.

Distribution of the Least Squares Estimators

Proof.

We will prove the statement for the slope only. Since

$$\sum_{i=1}^{n}(x_i - \overline{x}) = 0$$

we may write

$$B_1 = \frac{1}{S_{xx}}\sum_{i=1}^{n}(x_i - \overline{x})(Y_i - \overline{Y}) = \frac{1}{S_{xx}}\sum_{i=1}^{n}(x_i - \overline{x})Y_i \tag{24.3}$$

Now B_1 is a linear combination of the i.i.d. normally distributed Y_i, so B_1 itself follows a normal distribution.

It remains to show that B_1 has mean β_1 and variance $\sigma^2 / \sum_{i=1}^{n}(x_i - \overline{x})^2$.

Distribution of the Least Squares Estimators

Proof (continued).

$$
\begin{aligned}
\mathrm{E}[B_1] &= \mathrm{E}\left[\sum_{i=1}^{n} \frac{x_i - \overline{x}}{S_{xx}} Y_i\right] = \sum_{i=1}^{n} \frac{x_i - \overline{x}}{S_{xx}} \mathrm{E}[Y_i] \\
&= \sum_{i=1}^{n} \frac{(x_i - \overline{x})(\beta_0 + \beta_1 x_i)}{\sum_{i=1}^{n} (x_i - \overline{x})^2} \\
&= \frac{\beta_0}{\sum_{i=1}^{n} (x_i - \overline{x})^2} \underbrace{\sum_{i=1}^{n} (x_i - \overline{x})}_{=0} + \beta_1 \underbrace{\frac{\sum_{i=1}^{n} x_i (x_i - \overline{x})}{\sum_{i=1}^{n} (x_i - \overline{x})^2}}_{=1} \\
&= \beta_1 .
\end{aligned}
$$

Distribution of Least Squares Estimators

Proof (continued).

Similarly,

$$\begin{aligned}
\operatorname{Var} B_1 &= \operatorname{Var}\Big(\sum_{i=1}^{n} \frac{x_i - \overline{x}}{S_{xx}} Y_i\Big) = \sum_{i=1}^{n} \frac{(x_i - \overline{x})^2}{S_{xx}^2} \operatorname{Var} Y_i \\
&= \frac{\sigma^2}{\Big(\sum_{j=1}^{n}(x_j - \overline{x})^2\Big)^2} \sum_{i=1}^{n}(x_i - \overline{x})^2 \\
&= \frac{\sigma^2}{\sum_{i=1}^{n}(x_i - \overline{x})^2}.
\end{aligned}$$

The proof of the corresponding statement for the estimator B_0 is completely analogous. $\square$

Least Squares Estimator for the Variance

The variance σ^2 of $Y \mid x$ is assumed to be the same for all values of x. To estimate it, we use the error sum of squares,

$$S^2 := \frac{\mathrm{SS_E}}{n-2} = \frac{1}{n-2}\sum_{i=1}^{n}(y_i - \widehat{\mu}_{Y|x_i})^2 \tag{24.4}$$

It turns out that this estimator is unbiased for σ^2 and in fact

$$(n-2)S^2/\sigma^2 = \frac{\mathrm{SS_E}}{\sigma^2}$$

follows a chi-squared distribution with $n-2$ degrees of freedom.

Furthermore, it can be shown that S^2 is independent of B_0 and B_1.

(Analogously to the statement that the sample mean is independent of the sample variance, which we proved using the Helmert transformation.)

Inferences on the Slope and the Intercept

Therefore,

$$\frac{(B_1 - \beta_1)/(\sigma/\sqrt{S_{xx}})}{\sqrt{(n-2)S^2/[\sigma^2(n-2)]}} = \frac{B_1 - \beta_1}{S/\sqrt{S_{xx}}}$$

follows a T-distribution with $n-2$ degrees of freedom.

The same is true for

$$\frac{(B_0 - \beta_0)/(\sigma\sqrt{\sum x_k^2}/\sqrt{nS_{xx}})}{\sqrt{(n-2)S^2/[\sigma^2(n-2)]}} = \frac{B_0 - \beta_0}{S\sqrt{\sum x_k^2}/\sqrt{nS_{xx}}}.$$

It follows immediately that we have $100(1-\alpha)\%$ confidence intervals

$$B_1 \pm t_{\alpha/2,n-2}\frac{S}{\sqrt{S_{xx}}}, \qquad B_0 \pm t_{\alpha/2,n-2}\frac{S\sqrt{\sum x_i^2}}{\sqrt{nS_{xx}}}$$

for β_1 and β_0, respectively.

Practical Calculations

In practice, to simplify calculations using a calculator, we may use the following relations:

$$S_{xx} = \sum_{i=1}^{n} (x_i - \overline{x})^2 = \sum_{i=1}^{n} x_i^2 - \frac{1}{n}\Big(\sum_{i=1}^{n} x_i\Big)^2,$$

$$S_{yy} = \sum_{i=1}^{n} (y_i - \overline{y})^2 = \sum_{i=1}^{n} y_i^2 - \frac{1}{n}\Big(\sum_{i=1}^{n} y_i\Big)^2,$$

$$S_{xy} = \sum_{i=1}^{n} (x_i - \overline{x})(y_i - \overline{y}) = \sum_{i=1}^{n} x_i y_i - \frac{1}{n}\Big(\sum_{i=1}^{n} x_i\Big)\Big(\sum_{i=1}^{n} y_i\Big)$$

and

$$b_0 = \overline{y} - b_1 \overline{x}, \qquad b_1 = \frac{S_{xy}}{S_{xx}}, \qquad \mathrm{SS_E} = S_{yy} - b_1 S_{xy}.$$

Confidence Intervals for Slope and Intercept

24.5. Example. We return to Example 24.1 of solvent evaporation in spray painting. Recall that we obtained the point estimate for the regression line

$$\widehat{\mu}_{Y|x} = 13.6 - 0.0795x.$$

Based on the previously calculated

$$n = 25, \qquad \sum x_i = 1312.9, \qquad \sum y_i = 235.70,$$
$$\sum x_i^2 = 76193.7, \qquad \sum y_i^2 = 2286.07, \qquad \sum x_i y_i = 11802.2$$

we obtain

$$S_{xx} = 7245.47, \qquad S_{yy} = 63.89, \qquad S_{xy} = -575.781$$

and

$$\mathrm{SS_E} = S_{yy} - b_1 S_{xy} = 18.13, \qquad s^2 = \mathrm{SS_E}/(n-2) = 0.79$$

Confidence Intervals for Slope and Intercept

A 95% confidence interval for the slope of the regression line is given by

$$\begin{aligned} b_1 \pm t_{0.025,23} s/\sqrt{S_{xx}} &= -0.0795 \pm \frac{2.0687 \cdot 0.888}{85.12} \\ &= -0.0795 \pm 0.0215 \end{aligned}$$

and a 95% confidence interval for the intercept is given by

$$b_0 \pm t_{0.025,23} s \sqrt{\sum x_i^2} / \sqrt{n S_{xx}} = 13.6 \pm 1.19$$

These confidence intervals can also be obtained with Mathematica:

```
model["ParameterConfidenceIntervals", ConfidenceLevel → 0.95]
{{12.41, 14.7927}, {-0.101047, -0.0578881}}
```

Tests for Regression Parameters

Of course, we may also perform hypothesis tests (Fisher or Neyman-Pearson) on the model parameters. For example, we may test

$$H_0\colon \beta_0 = \beta_0^0 \qquad \text{and} \qquad H_0\colon \beta_1 = \beta_1^0$$

for null values β_0^0 and β_1^0 of the intercept and slope, respectively.

An important special case is following:medskip
We say that a regression is ***significant*** if there is statistical evidence that the slope $\beta_1 \neq 0$.

Test for Significance of Regression

24.6. Test for Significance of Regression. Let $(x_i, Y \mid x_i)$, $i = 1, \dots, n$ be a random sample from $Y \mid x$. We reject

$$H_0\colon \beta_1 = 0$$

at significance level α if the statistic

$$T_{n-2} = \frac{B_1}{S/\sqrt{S_{xx}}}.$$

satisfies $|T_{n-2}| > t_{\alpha/2,n-2}$.

Significance of Regression

24.7. Example. We return to Example 24.1 of solvent evaporation in spray painting. Recall that we obtained the point estimate for the regression line

$$\widehat{\mu}_{Y|x} = 13.64 - 0.08x.$$

We now test whether the regression is significant. from the previously calculated data, we have

$$t_{23} = \frac{b_1}{s/\sqrt{S_{xx}}} = -7.62.$$

We find that $P[T_{23} \le -7.62] < 0.0005$.

Since this is a two-tailed test, $P < 2 \cdot 0.0005 = 0.001$.

Hence, we are able to reject H_0. There is no evidence that the regression is not significant.

Properties of the Estimator for the Mean

We now turn to the actual estimated mean, $\mu_{Y|x}$. The least-squares estimators give

$$\widehat{\mu}_{Y|x} = B_0 + B_1 x.$$

Since B_0 and B_1 are unbiased estimators for β_0 and β_1 it follows immediately that $\widehat{\mu}_{Y|x}$ is unbiased for $\mu_{Y|x}$.

We may write

$$\widehat{\mu}_{Y|x} = B_0 + B_1 x = \overline{Y} - B_1 \overline{x} + B_1 x = \overline{Y} + B_1 (x - \overline{x}).$$

Since B_1 is a linear combination of the Y_i (see (24.3)), this implies that $\widehat{\mu}_{Y|x}$ is also a linear combination of the Y_i. The Y_i are assumed independent and normally distributed, so we see that $\widehat{\mu}_{Y|x}$ follows a normal distribution.

Distribution of the Estimated Mean

Since $\operatorname{Cov}(\overline{Y}, B_1) = 0$ (see assignments),

$$\operatorname{Var}[\widehat{\mu}_{Y|x}] = \operatorname{Var}[\overline{Y}] + (x - \overline{x})^2 \operatorname{Var} B_1 = \frac{\sigma^2}{n} + \frac{(x - \overline{x})^2 \sigma^2}{S_{xx}}.$$

In conclusion,

$$\frac{\widehat{\mu}_{Y|x} - \mu_{Y|x}}{\sigma \sqrt{\frac{1}{n} + \frac{(x - \overline{x})^2}{S_{xx}}}}$$

follows a standard-normal distribution.

Confidence Interval the Estimated Mean

Using our estimator for the variance, we see that

$$\frac{1}{\sqrt{(n-2)S^2/\sigma^2}/\sqrt{n-2}} \frac{\widehat{\mu}_{Y|x} - \mu_{Y|x}}{\sigma\sqrt{\frac{1}{n} + \frac{(x-\overline{x})^2}{S_{xx}}}} = \frac{\widehat{\mu}_{Y|x} - \mu_{Y|x}}{S\sqrt{\frac{1}{n} + \frac{(x-\overline{x})^2}{S_{xx}}}}$$

follows a T distribution with $n-2$ degrees of freedom.

Based on this, we may make inferences on the value of the mean of $Y \mid x$.

For example, we obtain the following $100(1-\alpha)\%$ confidence interval for $\mu_{Y|x}$:

$$\widehat{\mu}_{Y|x} \pm t_{\alpha/2,n-2} S\sqrt{\frac{1}{n} + \frac{(x-\overline{x})^2}{S_{xx}}}. \tag{24.5}$$

Inferences about a Single Predicted Value

We are interested in finding an "estimate" (guess) or a ***prediction*** for the value of the random variable $Y \mid x$. Note the essential difference:

- An ***estimate*** is a statistical statement on the value of an unknown, but fixed, population parameter.
- A ***prediction*** is a statistical statement on the value of an essentially random quantity.

We define a $100(1-\alpha)\%$ prediction interval $[L_1, L_2]$ for a random variable X by

$$P[L_1 \le X \le L_2] = 1 - \alpha.$$

As a ***predictor*** $\widehat{Y \mid x}$ for the value of $Y \mid x$ we use the estimator for the mean, i.e., we set

$$\widehat{Y \mid x} = \widehat{\mu}_{Y|x} = B_0 + B_1 x.$$

In order to find a prediction interval, we need to analyze the distribution of $\widehat{Y \mid x}$.

Inferences about a Single Predicted Value

Recall that $\widehat{\mu}_{Y|x}$ follows a normal distribution with mean $\mu_{Y|x}$ and variance $\left(\frac{1}{n} + \frac{(x-\overline{x})^2}{S_{xx}}\right)\sigma^2$. Furthermore, $Y \mid x$ is normally distributed with mean $\mu_{Y|x}$ and variance σ^2.

Hence $\widehat{Y \mid x} - Y \mid x$ is normally distributed and, furthermore,

$$\begin{aligned} \mathrm{E}[\widehat{Y \mid x} - Y \mid x] &= \mu_{Y|x} - \mu_{Y|x} = 0, \\ \mathrm{Var}[\widehat{Y \mid x} - Y \mid x] &= \left(\frac{1}{n} + \frac{(x-\overline{x})^2}{S_{xx}}\right)\sigma^2 + \sigma^2 = \left(1 + \frac{1}{n} + \frac{(x-\overline{x})^2}{S_{xx}}\right)\sigma^2. \end{aligned}$$

Thus, after standardizing and dividing by S/σ we obtain the T_{n-2} random variable

$$T_{n-2} = \frac{\widehat{Y \mid x} - Y \mid x}{S\sqrt{1 + \frac{1}{n} + \frac{(x-\overline{x})^2}{S_{xx}}}}$$

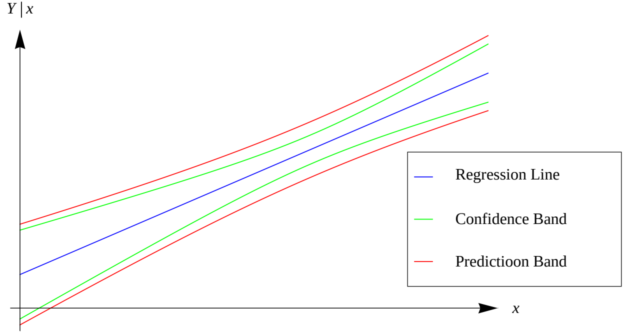
JOINT INSTITUTE 交大密西根学院

Inferences about a Single Predicted Value

We thus obtain the following $100(1-\alpha)\%$ prediction interval for $Y \mid x$:

$$\widehat{Y \mid x} \pm t_{\alpha/2,n-2} S \sqrt{1 + \frac{1}{n} + \frac{(x-\overline{x})^2}{S_{xx}}} \tag{24.6}$$

The limits of the confidence interval (24.5) and the prediction interval (24.6), plotted as functions of x, are commonly called ***confidence bands*** and ***prediction bands*** for the regression.



Confidence and Prediction Intervals

24.8. Example. Continuing with the data from Example 24.1, Mathematica gives confidence bands (24.5) for the estimated mean as

```
conf = model["MeanPredictionBands", ConfidenceLevel → 0.95]
```

$$\left\{13.6013 - 0.0794677\,x - 2.06866\sqrt{0.331656 - 0.0114296\,x + 0.00010882\,x^2},\right.$$
$$\left.13.6013 - 0.0794677\,x + 2.06866\sqrt{0.331656 - 0.0114296\,x + 0.00010882\,x^2}\right\}$$

Prediction bands (24.6) are given by

```
pred = model["SinglePredictionBands", ConfidenceLevel → 0.95]
```

$$\left\{13.6013 - 0.0794677\,x - 2.06866\sqrt{1.12011 - 0.0114296\,x + 0.00010882\,x^2},\right.$$
$$\left.13.6013 - 0.0794677\,x + 2.06866\sqrt{1.12011 - 0.0114296\,x + 0.00010882\,x^2}\right\}$$

Confidence and Prediction Intervals

Below, the prediction bands are shown in red, while the confidence bands for the estimated mean are green: