Statistics

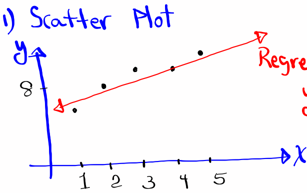
Lecture 16

Feb 19-8:47 AM

SG 33

x	y
2	8
1	6
3	9
4	9
5	10

1) Scatter Plot



Regression line

$\hat{y} = a + bx$

Clear all lists

$x \to L1$, $y \to L2$

STAT CALC LinReg (a+bx)

Xlist: L1

Ylist: L2

FreqList

$a = 5.7$

$b = .9$

$\to$ $\hat{y} = 5.7 + .9x$

$r^2 = .880$ $r^2(\%) = 88\%$

$r = .938$

Linear Correlation Coef.

Coef. of determination

88% of Y-values are explained by x-values.

If r is close to 1 or −1, Linear Correlation is significant

If r is close to 0, Linear Correlation is not significant

Dec 12-7:33 PM

Testing Linear Correlation:

$H_0: \rho = 0$ ← rho — Linear correlation is not significant

$H_1: \rho \neq 0$ TTT — Linear Correlation is Significant.

STAT → TESTS ↓ LinRegTTest

xlist: L1

Ylist: L2

Freq: 1

$\rho \neq 0$

RegEQ:

Calculate

CTS $t = 4.700$

P-value $P = .018$

Suppose $\alpha = .1$ and we want to determine whether r is significant or not.

P-value $\leq \alpha$
.018 ≤ .1

H_0 invalid

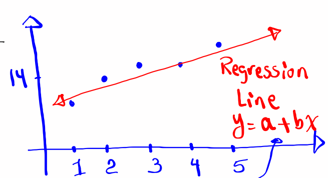
H_1 valid → Linear Correlation is Significant

$H_0: \rho = 0$ Not Significant

$H_1: \rho \neq 0$ Is Significant

Dec 12-7:42 PM

Study time	QZ Scores
2	14
3	15
4	15
1	12
5	18



Study time → x → L1

QZ Score → y → L2

STAT → CALC ↓ LinReg(a+bx)

$a = 10.9$

$b = 1.3$

$r^2 = .899 \approx 90\%$

$r = .948$

$y = 10.9 + 1.3x$

Coef. of determination $r^2 \approx 90\%$ — 90% of QZ Scores are explained by Study time.

Linear Correlation Coef. $r = .948$

Dec 12-7:51 PM

Use $\alpha = .02$ to test the claim that the Linear Correlation is Significant.

$H_0: \rho = 0$ Not Significant

$H_1: \rho \neq 0$ Is Significant ↑↑↑ claim

STAT → TESTS ↓ LinRegTTest

Xlist: L1
Ylist: L2
Freq: 1
$\rho \neq 0$
Calculate

CTS $t = 5.166$
P-Value $p = .014$

$df = 3$

$df = n-2$

P-value $\leq \alpha$
.014 .02

H_0 invalid

H_1 valid Linear Correlation is Significant.

Dec 12-7:58 PM

Formula for CTS t

$r = .948$
$r^2 = .899$
$n = 5$

$$t = r \cdot \sqrt{\frac{n-2}{1-r^2}}$$

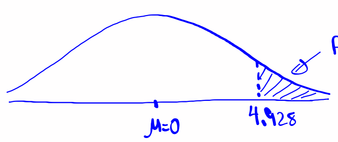
$$= .948 \cdot \sqrt{\frac{5-2}{1-.899}} = .948 \cdot \sqrt{\frac{3}{.111}}$$

$$\approx 4.928$$

STAT TESTS

LinRegTTest

CTS $t =$
P-Value $P =$



Area .2 = P-Value

$2 \cdot tcdf(4.928, E99, 3)$

$= .016$

σ unknown
$df = 3$

P-Value $\leq \alpha$
.016 .02

H_0 invalid

H_1 valid → Linear Correlation is Significant.

Dec 12-8:05 PM

Suppose $n = 10$, $r = -.857$, $r^2 = .734$, $\alpha = .1$

$df = n - 2 = 8$

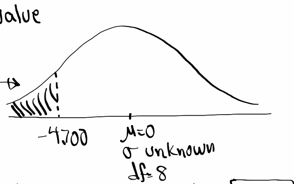
CTS $t = r \cdot \sqrt{\frac{n-2}{1-r^2}}$

$= -.857 \cdot \sqrt{\frac{10-2}{1-.734}} = -.857 \cdot \sqrt{\frac{8}{.266}}$

$\approx \boxed{-4.700}$

Find P-value

P-value $= 2 \cdot$ Area



$= 2 \cdot \text{tcdf}(-E99, -4.700, 8) = \boxed{.002}$

If $\alpha = .1$
P-value $\le \alpha$ → H_0 invalid, H_1 Valid → Linear Correlation is Significant.

Dec 12-8:16 PM

How to make predictions

If r is Significant ⇒ Use the regression line
Plug in x, find y

If r is not Significant ⇒ Use $\bar{y}$

Dec 12-8:24 PM

Suppose $\hat{y} = 8 + 2.5x$, $\bar{y} = 15$ $r = .6$

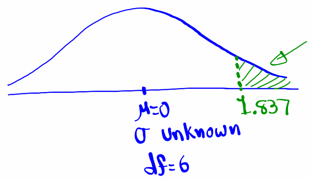
$n = 8$ Predict y if $x = 4$. → $df = n - 2 = 6$

First test Linear correlation r

$H_0: \rho = 0$ Not Significant

$H_1: \rho \neq 0$ Is Significant

CTS $t = r \cdot \sqrt{\frac{n-2}{1-r^2}} = .6 \cdot \sqrt{\frac{8-2}{1-.36}} = .6 \cdot \sqrt{\frac{6}{.64}} = \boxed{1.837}$



Area · 2 = P-value

P-value = 2 · tcdf(1.837, E99, 6)

= $\boxed{.116}$

No α given ⇒ Use .05

P-value > α
.116 .05

H_0 Valid → Linear Correlation
H_1 Invalid is not Significant
use $\bar{y}$
(15)

Dec 12-8:26 PM

Given $\hat{y} = 25 - 4x$, $\bar{y} = 10$, $r = -.7$ $n = 12$

$\alpha = .1$ Predict y when $x = 3$.

If r is significant → $\hat{y} = 25 - 4(3) = \boxed{13}$

If r is not Significant → $\bar{y} = \boxed{10}$

$H_0: \rho = 0$ Not Significant

$H_1: \rho \neq 0$ Is Significant (TTT)

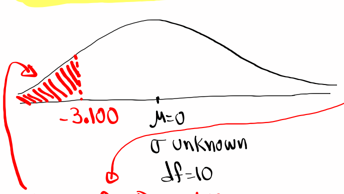
CTS t

$t = r \cdot \sqrt{\frac{n-2}{1-r^2}}$

$= -.7 \cdot \sqrt{\frac{12-2}{1-.49}}$

$= -.7 \cdot \sqrt{\frac{10}{.51}}$

$= \boxed{-3.100}$



Area · 2 = P-value

P-value = 2 · tcdf(−E99, −3.100, 10) = $\boxed{.011}$ ✓

P-value ≤ α → H_0 invalid
.011 .1

H_1 Valid → Linear Correlation is Significant

⇒ $\hat{y} = 25 - 4(3) = \boxed{13}$

SG 33 ✓

Dec 12-8:36 PM