**Q22**

(a)

When the Species is Hemigrapsus nudus

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 0.5191 1.1147 0.466 0.650

log(Height) 0.4083 0.5426 0.752 0.466

**So, the estimated slope in the SLR of log force on log height is 0.4083. The standard error of the estimated slope in the simple linear regression of log force on log height for Hemigrapsus nudus is****0.5426**

When the Species is Lophopanopeus bellus

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) -3.7800 1.2842 -2.943 0.014699 \*

log(Height) 2.9737 0.6125 4.855 0.000666 \*\*\*

**So, the estimated slope in the SLR of log force on log height is 2.9737. The standard error of the estimated slope in the simple linear regression of log force on log height for Lophopanopeus bellus is** **0.6125**

When the Species is Cancer productus

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) -1.9673 0.9978 -1.972 0.076950 .

log(Height) 2.0685 0.4275 4.838 0.000683 \*\*\*

**So, the estimated slope in the SLR of log force on log height is 2.0685. The standard error of the estimated slope in the simple linear regression of log force on log height for** **Cancer productus is** **0.4275**

(b)

Consider the C. productus and L. bellus

SE[beta-hat1(1)-beta-hat1(2)]=( 0.4275 ^2+0.6125^2)^0.5=0.7469354

Consider the C. productus and H. nudus

SE[beta-hat1(1)-beta-hat1(2)]=( 0.4275 ^2+ 0.5426^2)^0.5=0.6907757

Thus, it can conclude that the C. productus and H. nudus are more consistent since the standard error is least and the C. productus and L. bellus has most points near about the linear regression since the standard deviation is the least in that case

**Q24**

(a)

Denmark:

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 5.987e-01 4.080e-02 14.673 <2e-16 \*\*\*

Year -4.289e-05 2.069e-05 -2.073 0.0442 \*

The Netherlands:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 6.724e-01 2.792e-02 24.08 < 2e-16 \*\*\*

Year -8.084e-05 1.416e-05 -5.71 9.64e-07 \*\*\*

Canada:

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 7.338e-01 5.480e-02 13.390 3.98e-11 \*\*\*

Year -1.112e-04 2.768e-05 -4.017 0.000738 \*\*\*

USA:

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 6.201e-01 1.860e-02 33.340 < 2e-16 \*\*\*

Year -5.429e-05 9.393e-06 -5.779 1.44e-05 \*\*\*

Thus, the data are all confirm with the presented one.

(b)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Beta1 | SE(Beta1) | T-statistic | p-value |
| Denmark | -0.00004289 | 0.00002069 | -2.073 | 0.044 |
| Netherlands | -0.00008084 | 0.00001416 | -5.710 | < 0.0005 |
| Canada | -0.0001112 | 0.00002768 | -4.017 | 0.001 |
| USA | -0.00005429 | 0.000009393 | -5.779 | < 0.0005 |

Since the p-value for Beta1 are all less than 0.05 in the four countries, thus we can conclude that the proportion of male births is truly declining.

(c)

Because USA has the smallest SE(Beta1) which means the observations have less variability among the estimated regression line and the true values are closely fitted by the line. Thus, while the beta1 is small, the t-statistic can be large.

(d)

Standard error measures the closeness of predicted data with the observed data. The regression line is more appropriately fit for USA rather than Canada. Hence, standard error of slope is lesser for USA than Canada.

(E)

The distribution changes, the standard deviation values will change as well.

It is probably caused by the difference of population among 4 countries. With more population, the birth rate of male will be closer to 50% and has a less fluctuation.

Q26

> ADJ=Height\*(Gender=="male") + 1.08\*Height\*(Gender=="female")

> MID=(Father+1.08\*Mother)/2

> m=lm(formula=ADJ~Height)

> summary(m)

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 19.82606 2.82206 7.025 4.12e-12 \*\*\*

MID 0.71392 0.04076 17.513 < 2e-16 \*\*\*

Thus, we can get

y-hat=0.71392X+19.82606

-parents are 65 inches on average

y-hat =19.82606+0.71392\*65=66.23086

-parents are 76 inches on average

y-hat=19.82606+0.71392\*76=74.08398

-parents are 65 inches on average

> height=data.frame(MID=65)

> predict(m,height,interval="prediction",level=0.95)

fit lwr upr

1 66.23091 61.81268 70.64915

Thus 95% prediction interval is (61.81268,70.64915)

-parents are 76 inches on average

> height=data.frame(MID=76)

> predict(m,height,interval="prediction",level=0.95)

fit lwr upr

1 74.08404 69.64523 78.52285

Thus 95% prediction interval is(69.64532,78.52285)