实验3.11

谭铭瑞

2023-04-14

library(haven)  
xt3\_11 <- read\_sav("G:\\Rwork\\test3\\xt3.11.sav")  
View(xt3\_11)  
data3\_11<-data.frame(xt3\_11)[-11,]  
data3\_11

## Y X1 X2 X3  
## 1 160 70 35 1.0  
## 2 260 75 40 2.4  
## 3 210 65 40 2.0  
## 4 265 74 42 3.0  
## 5 240 72 38 1.2  
## 6 220 68 45 1.5  
## 7 275 78 42 4.0  
## 8 160 66 36 2.0  
## 9 275 70 44 3.2  
## 10 250 65 42 3.0

cor(data3\_11)#相关系数矩阵

## Y X1 X2 X3  
## Y 1.0000000 0.5556527 0.7306199 0.7235354  
## X1 0.5556527 1.0000000 0.1129513 0.3983870  
## X2 0.7306199 0.1129513 1.0000000 0.5474739  
## X3 0.7235354 0.3983870 0.5474739 1.0000000

lm311<-lm(Y~X1+X2+X3,data = data3\_11)#建y~x1+x2+x3立回归方程  
summary(lm311)

##   
## Call:  
## lm(formula = Y ~ X1 + X2 + X3, data = data3\_11)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -25.198 -17.035 2.627 11.677 33.225   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -348.280 176.459 -1.974 0.0959 .  
## X1 3.754 1.933 1.942 0.1002   
## X2 7.101 2.880 2.465 0.0488 \*  
## X3 12.447 10.569 1.178 0.2835   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 23.44 on 6 degrees of freedom  
## Multiple R-squared: 0.8055, Adjusted R-squared: 0.7083   
## F-statistic: 8.283 on 3 and 6 DF, p-value: 0.01487

lm2<-lm(Y~X1+X2,data = data3\_11)#剔除X3后重新建立回归方程  
summary(lm2)

##   
## Call:  
## lm(formula = Y ~ X1 + X2, data = data3\_11)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -42.012 -10.656 4.358 11.984 28.927   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -459.624 153.058 -3.003 0.01986 \*   
## X1 4.676 1.816 2.575 0.03676 \*   
## X2 8.971 2.468 3.634 0.00835 \*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 24.08 on 7 degrees of freedom  
## Multiple R-squared: 0.7605, Adjusted R-squared: 0.6921   
## F-statistic: 11.12 on 2 and 7 DF, p-value: 0.006718

confint(lm2)#置信水平为95%的置信区间

## 2.5 % 97.5 %  
## (Intercept) -821.5473012 -97.700006  
## X1 0.3813047 8.969956  
## X2 3.1339785 14.807944

#标准化  
data311\_standard<-scale(data3\_11)/sqrt(nrow(data3\_11)-1)  
data311\_new<-data.frame(data311\_standard)  
data311\_new#样本相关矩阵

## Y X1 X2 X3  
## 1 -0.54914770 -0.02247964 -0.54999057 -0.45886628  
## 2 0.21889104 0.35218098 -0.04074004 0.02415086  
## 3 -0.16512833 -0.39714025 -0.04074004 -0.11385404  
## 4 0.25729298 0.27724886 0.16296017 0.23115820  
## 5 0.06528329 0.12738461 -0.24444025 -0.38986384  
## 6 -0.08832445 -0.17234388 0.46851049 -0.28636016  
## 7 0.33409685 0.57697735 0.16296017 0.57617045  
## 8 -0.54914770 -0.32220813 -0.44814046 -0.11385404  
## 9 0.33409685 -0.02247964 0.36666038 0.30016065  
## 10 0.14208717 -0.39714025 0.16296017 0.23115820

lm3<-lm(Y~X1+X2+X3,data = data311\_new)  
summary(lm3)

##   
## Call:  
## lm(formula = Y ~ X1 + X2 + X3, data = data311\_new)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -0.19353 -0.13083 0.02017 0.08969 0.25518   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -1.277e-16 5.693e-02 0.000 1.0000   
## X1 3.848e-01 1.982e-01 1.942 0.1002   
## X2 5.355e-01 2.172e-01 2.465 0.0488 \*  
## X3 2.771e-01 2.353e-01 1.178 0.2835   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.18 on 6 degrees of freedom  
## Multiple R-squared: 0.8055, Adjusted R-squared: 0.7083   
## F-statistic: 8.283 on 3 and 6 DF, p-value: 0.01487

newdata<-data.frame(75,42,3.1)  
names(newdata)<-c("X1","X2","X3")  
pred\_y<-predict(lm2,newdata)  
pred\_y

## 1   
## 267.829

pred\_yplim<-predict(lm2,newdata,interval = "prediction")#计算预测区间  
pred\_yplim

## fit lwr upr  
## 1 267.829 204.4355 331.2225

pred\_yclim<-predict(lm2,newdata,interval = "confidence")#计算置信区间  
pred\_yclim

## fit lwr upr  
## 1 267.829 239.9677 295.6903

#基本分析：  
#货运量与工业总产值与农业总产值有较大相关性  
#居民非商品支出与货运总量的相关性不大，p值>0.05被优先从模型中剔除

实验3.12

谭铭瑞

2023-04-14

library(haven)  
xt3\_12 <- read\_sav("G:\\Rwork\\test3\\xt3.12 三次产业.sav")  
View(xt3\_12)  
data3\_12<-data.frame(xt3\_12)  
data3\_12

## year GDP x1 x2 x3  
## 1 1990 18667.8 5062.0 7717.4 5888.4  
## 2 1991 21781.5 5342.2 9102.2 7337.1  
## 3 1992 26923.5 5866.6 11699.5 9357.4  
## 4 1993 35333.9 6963.8 16454.4 11915.7  
## 5 1994 48197.9 9572.7 22445.4 16179.8  
## 6 1995 60793.7 12135.8 28679.5 19978.5  
## 7 1996 71176.6 14015.4 33835.0 23326.2  
## 8 1997 78973.0 14441.9 37543.0 26988.1  
## 9 1998 84402.3 14817.6 39004.2 30580.5  
## 10 1999 89677.1 14770.0 41033.6 33873.4  
## 11 2000 99214.6 14944.7 45555.9 38714.0  
## 12 2001 109655.2 15781.3 49512.3 44361.6  
## 13 2002 120332.7 16537.0 53896.8 49898.9  
## 14 2003 135822.8 17381.7 62436.3 56004.7  
## 15 2004 159878.3 21412.7 73904.3 64561.3  
## 16 2005 184937.4 22420.0 87598.1 74919.3  
## 17 2006 216314.4 24040.0 103719.5 88554.9  
## 18 2007 265810.3 28627.0 125831.4 111351.9  
## 19 2008 314045.4 33702.0 149003.4 131340.0  
## 20 2009 340902.8 35226.0 157638.8 148038.0  
## 21 2010 401512.8 40533.6 187383.2 173596.0  
## 22 2011 473104.0 47486.2 220412.8 205205.0  
## 23 2012 518942.1 52373.6 235162.0 231406.5

cor(data3\_12)#相关系数矩阵

## year GDP x1 x2 x3  
## year 1.0000000 0.9231405 0.9348400 0.9262837 0.9157724  
## GDP 0.9231405 1.0000000 0.9953124 0.9996337 0.9994158  
## x1 0.9348400 0.9953124 1.0000000 0.9949552 0.9928450  
## x2 0.9262837 0.9996337 0.9949552 1.0000000 0.9982887  
## x3 0.9157724 0.9994158 0.9928450 0.9982887 1.0000000

lm312<-lm(GDP~x1+x2+x3,data = data3\_12)#建y~x1+x2立回归方程  
summary(lm312)

##   
## Call:  
## lm(formula = GDP ~ x1 + x2 + x3, data = data3\_12)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -0.102845 -0.007457 -0.004082 -0.000857 0.094929   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 8.079e-03 3.796e-02 2.130e-01 0.834   
## x1 1.000e+00 6.273e-06 1.594e+05 <2e-16 \*\*\*  
## x2 1.000e+00 2.480e-06 4.032e+05 <2e-16 \*\*\*  
## x3 1.000e+00 2.161e-06 4.627e+05 <2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.03919 on 19 degrees of freedom  
## Multiple R-squared: 1, Adjusted R-squared: 1   
## F-statistic: 1.049e+14 on 3 and 19 DF, p-value: < 2.2e-16

#x1,x2,x3的p值远小于0.001因此我们可以拒绝原假设，认为自变量x1,x2,x3 对因变量 y 有显著影响