punto14

April 8, 2022

[ ]:
```
GPA<-c(2.8,3.4,3,3.5,3.6,3,2.7,3.7)
ACT<-c(21,24,26,27,29,25,25,30)
df<-data.frame(ACT,GPA)
df
```

A data.frame: 8 × 2

ACT <dbl>	GPA <dbl>
21	2.8
24	3.4
26	3.0
27	3.5
29	3.6
25	3.0
25	2.7
30	3.7

[ ]:
```
X<-data.frame(1.0,df$ACT)
Y<-data.frame(df$GPA)
Y<-data.matrix(Y)
X<-data.matrix(X)
X_t<-t(X)
X_tX<-X_t %*% X
X_tX_inv<-solve(X_tX)
X_tY<-X_t %*% Y

BETA<-X_tX_inv %*% X_tY


BETA
```

A matrix: 2 × 1 of type dbl

	df.GPA
X1	0.5681319
df.ACT	0.1021978

[ ]:
```
cat('if ACT up in 7.0, GPA will be up in ',BETA[2,1]*7,'\n')
```

```
if ACT up in 7.0, GPA will be up in  0.7153846
```

[ ]:
```
df$GPA_pred<-X %*% BETA
df$ERRORS<-df$GPA-df$GPA_pred

df
```

A data.frame: 8 × 4

ACT <dbl>	GPA <dbl>	GPA_pred <dbl[,1]>	ERRORS <dbl[,1]>
21	2.8	2.714286	0.08571429
24	3.4	3.020879	0.37912088
26	3.0	3.225275	-0.22527473
27	3.5	3.327473	0.17252747
29	3.6	3.531868	0.06813187
25	3.0	3.123077	-0.12307692
25	2.7	3.123077	-0.42307692
30	3.7	3.634066	0.06593407

[ ]:
```
cat('the sum of errors is ',sum(round(df$ERRORS,0)))
```

```
the sum of errors is  0
```

[ ]:
```
cat('if ACT is 20, GPA will be ',BETA[1,1]+BETA[2,1]*20)
```

```
if ACT is 20, GPA will be  2.612088
```

[ ]:
```
var_gpa<-var(df$GPA)
var_gpa_prec<-var(df$GPA_pred)
var_gpa_prec<-var_gpa_prec[1,1]

cat('the variance of GPA is ',var_gpa,'\n')
cat('the variance of GPA_pred is ',var_gpa_prec,'\n')

cat('R^2 is ',var_gpa_prec/var_gpa,'\n')
```

```
the variance of GPA is  0.1469643
the variance of GPA_pred is  0.08486068
R^2 is  0.5774238
```

una gran proporcion de la varianza de GPA esta explicada por ACT, por lo que se puede concluir que ACT esta siendo un buen predictor de GPA