

```

carpet <- read.csv("http://www.stat.ufl.edu/~winner/data/carpet_age.csv",
  header=T)
attach(carpet); names(carpet)

### f ==> full data m==> missing age
age_f <- age[1:23]; age_m <- age[24:25]
cys_acid_f <- cys_acid[1:23]; cys_acid_m <- cys_acid[24:25]

##### lm function
carpet.mod1 <- lm(age_f ~ cys_acid_f)
summary(carpet.mod1)
anova(carpet.mod1)
confint(carpet.mod1)
predict(carpet.mod1,list(cys_acid_f=cys_acid_m),int="p")

plot(cys_acid_f, age_f)
abline(carpet.mod1)

cor.test(cys_acid_f, age_f)

##### Scalar Form - Direct Calculations
(n <- length(age_f))
(ybar <- mean(age_f))
(xbar <- mean(cys_acid_f))
(SS_XX <- sum((cys_acid_f - xbar)^2))
(SS_XY <- sum((cys_acid_f - xbar) * (age_f - ybar)))
(SS_YY <- sum((age_f - ybar)^2))

(beta1_hat <- SS_XY / SS_XX)
(beta0_hat <- ybar - beta1_hat * xbar)
Y_hat <- beta0_hat + beta1_hat * cys_acid_f

(SS_ERR <- sum((age_f - Y_hat)^2)); (df_ERR <- n-2); (MS_ERR <- SS_ERR/df_ERR)
(SS_REG <- sum((Y_hat - ybar)^2)); (df_REG <- 1); (MS_REG <- SS_REG/df_REG)

(SE_beta1_hat <- sqrt(MS_ERR / SS_XX))
(SE_beta0_hat <- sqrt(MS_ERR * ((1/n) + xbar^2/SS_XX)))

(t_beta1 <- beta1_hat / SE_beta1_hat)
(t_beta0 <- beta0_hat / SE_beta0_hat)

(t_crit <- qt(.975,n-2))

(P_beta1 <- 2*(1-pt(abs(t_beta1),n-2)))
(P_beta0 <- 2*(1-pt(abs(t_beta0),n-2)))

(CI95_beta1 <- beta1_hat + qt(c(.025,.975),n-2) * SE_beta1_hat)
(CI95_beta0 <- beta0_hat + qt(c(.025,.975),n-2) * SE_beta0_hat)

X_s <- seq(0,5,0.01)
yhat_h <- beta0_hat + beta1_hat * X_s
CI_LO <- yhat_h + qt(.025,n-2) * sqrt(MS_ERR*((1/n)+(X_s-xbar)^2/SS_XX))
CI_HI <- yhat_h + qt(.975,n-2) * sqrt(MS_ERR*((1/n)+(X_s-xbar)^2/SS_XX))
PI_LO <- yhat_h + qt(.025,n-2) * sqrt(MS_ERR*(1 + (1/n)+(X_s-xbar)^2/SS_XX))
PI_HI <- yhat_h + qt(.975,n-2) * sqrt(MS_ERR*(1 + (1/n)+(X_s-xbar)^2/SS_XX))

plot(cys_acid_f,age_f,xlim=c(0,5),ylim=c(-500,2500))
lines(X_s,yhat_h,lty=1)
lines(X_h,CI_LO,lty=2)
lines(X_h,CI_HI,lty=2)
lines(X_h,PI_LO,lty=5)
lines(X_h,PI_HI,lty=5)

(yhat_miss <- beta0_hat + beta1_hat * cys_acid_m)

(PE_miss <- sqrt(MS_ERR * (1 + (1/n) + (cys_acid_m - xbar)^2/SS_XX)))

(PI_age_24 <- yhat_miss[1] + qt(c(.025,.975),n-2) * PE_miss[1])
(PI_age_25 <- yhat_miss[2] + qt(c(.025,.975),n-2) * PE_miss[2])

```

```
(F_obs <- MS_REG / MS_ERR)
(F_crit <- qf(.95,1,n-2))
(P_F <- 1 - pf(F_obs,1,n-2))

(r_square <- SS_REG / SS_YY)

### Correlation Test/CI

(r <- SS_XY / sqrt(SS_XX * SS_YY))
(t_r <- sqrt(n-2) * (r / sqrt(1 - r^2)))
(t_r_crit <- qt(.975,n-2))
(P_r <- 2*(1 - pt(abs(t_r),n-2)))

(z_r <- 0.5 * log((1+r) / (1-r)))
(CI_z_rho <- z_r + qnorm(c(.025,.975),0,1) * sqrt(1 / (n-3)))
(CI_rho <- (exp(2*CI_z_rho) - 1) / (exp(2*CI_z_rho) + 1))
```