Calculation of marginal probability distributions and probabilities:

$f_X(x_1) = \sum_y f_{X,Y}(x_1, y) = f_{X,Y}(x_1, y_1) + f_{X,Y}(x_1, y_2) + \cdots [or]\ \ f_X(x_1) = \int_y\ f_{X,Y}(x_1, y)\, dy$

independence: $For\ A, B: when\ P(A|B) = P(A)\ and\ P(B|A) = P(B) \qquad \left[P(A|B) = \frac{P(A \cap B)}{P(B)}\right]$

$Cov(X,Y) = E\big((X - \mu_X)(Y - \mu_Y)\big) \qquad Cov > 0: on\ average\ X, Y\ above\ (below) mean\ together.$

$Cov < 0: on\ average, when\ X\ is\ above\ (below)\ mean, Y\ is\ below\ (above).$

$(X,Y): independent \Longrightarrow Cov(X,Y) = 0 \qquad Cov(X,Y) = 0 \nRightarrow (X,Y): independent$

$Cov\big((a_1X + b_1), (a_2Y + b_2)\big) = a_1a_2Cov(X,Y) \quad Var(aX + bY) = a^2Var(X) + b^2Var(Y) + 2abCov(X,Y)$

$Corr(X,Y) = \rho_{XY} = \frac{Cov(X,Y)}{\sigma_X\sigma_Y} \qquad -1 \le \rho \le 1$

$(X,Y): independent \Longrightarrow Corr(X,Y) = 0 \qquad Corr(X,Y) = 0 \nRightarrow (X,Y): independent$

$E(X) = \sum_{i=1}^{k} x_i f_i(x_i)\ \ [or] \int_{-\infty}^{\infty} xf(x)dx \qquad E\big(g(x)\big) = \sum_{i=1}^{k} g(x_i) f_i(x_i)\ \ [or]\ \ \int_{-\infty}^{\infty} g(x)xf(x)dx$

$E(c) = c, \quad E(aX + b) = aE(X) + b, \quad E(X + Y) = E(X) + E(Y)$

$Var(X) = E(X - \mu)^2 = \sigma^2 = E(X^2) - \big(E(X)\big)^2 \qquad Var(c) = 0, \quad Var(aX + b) = a^2Var(X)$

$Z = \frac{X - \mu}{\sigma} \sim N(0,1) \qquad \Phi(\mathrm{c}) = z - table\ reading\ for\ number\ c$

$P(Z > z) = 1 - \Phi(\mathrm{z}), \qquad P(Z < -z) = P(Z > z), \qquad P(a \le Z \le b) = \Phi(\mathrm{b}) - \Phi(\mathrm{a})$

$P(|Z| > c) = P(Z > c) + P(Z < -c) = 2\big(1 - \Phi(\mathrm{c})\big)$

$E(Y|X = x) = \sum_j y_j f_{Y|X}(y_i|x) \quad or \quad \int y_i f_{Y|X}(y_i|x)dy$

Simple linear regression model. $y = \beta_0 + \beta_1 x + u$ assuming $E\langle u|x\rangle = 0 \Longrightarrow E\langle y|x\rangle = \beta_0 +$

$\beta_1 x \quad E(u) = 0 \Longrightarrow E(y - \beta_0 - \beta_1 x) = 0 \quad Cov(x,u) = E(xu) = 0 \Longrightarrow E\big(x(y - \beta_0 - \beta_1 x)\big) = 0$

$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1\bar{x} \qquad \hat{\beta}_1 = \frac{\sum_{i=1}^{n} x_i(y_i - \bar{y})}{\sum_{i=1}^{n} x_i(x_i - \bar{x})} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{n}(x_i - \bar{x})^2} \qquad \hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x \qquad \hat{u}_i = y_i - \hat{y}_i$

$\sum_{i=1}^{n} \hat{u}_i = 0 \qquad \sum_{i=1}^{n} x_i\hat{u}_i = 0 \qquad The\ point\ (\bar{x}, \bar{y}) is\ always\ on\ the\ regression\ line.$

$bias\big(\widehat{\beta_1}\big) = E\big(\widehat{\beta_1}\big) - \beta_1 \qquad E\big(\widehat{\beta_1}\big) = \beta_1 + E\left(\frac{\sum u_i x_i}{\sum x_i^2}\right) = \beta_1 + \frac{cov(u,x)}{var(x)} = \beta_1 + 0$

$\therefore bias\big(\widehat{\beta_1}\big) = \beta_1 - \beta_1 = 0 \ \blacksquare$

$SST = total\ sum\ of\ squares = total\ variation\ of\ y\ about\ its\ mean\ \ \sum(y_i - \bar{y})^2$

$SST_X = total\ variation\ of\ x\ about\ its\ mean\ \ \sum(x_i - \bar{x})^2 \qquad SSR = \sum(y_i - \hat{y}_i)^2$

$SSE = explained\ sum\ of\ squares = predicted\ variation\ of\ y\ about\ \bar{y} = \sum(\hat{y}_i - \bar{y})^2$

$R^2 = \frac{SSE}{SST} = 1 - \frac{SSR}{SST} \qquad shows\ fraction\ of\ variation\ in\ y\ explained\ by\ x$

$E\big(\hat{\beta}_0\big) = \beta_0 \qquad E\big(\hat{\beta}_1\big) = \beta_1 \qquad Var\big(\hat{\beta}_0\big) = \frac{\sigma^2 \sum x_i^2}{nSST_x} \qquad Var\big(\hat{\beta}_1\big) = \frac{\sigma^2}{SST_x}$

$[SST_x = \sum(x_i - \bar{x})^2]\ \ Unbiased\ estimator\ of\ population\ variance: \hat{\sigma}^2 = \frac{1}{n-2}\sum\hat{u}_i^2 = \frac{SSR}{n-2}$

$SER = \hat{\sigma} = \sqrt{\hat{\sigma}^2}$ $SE(\hat{\beta}_0) = \sqrt{Var(\hat{\beta}_0)} = \sqrt{\frac{\sigma^2 \sum x_i^2}{nSST_x}}$ $SE(\hat{\beta}_1) = \sqrt{Var(\hat{\beta}_1)} = \sqrt{\frac{\sigma^2}{SST_x}}$

$modify\ dependent$: $cy = c\beta_0 + c\beta_1 x$ $modify\ independent$: $y = \beta_0 + \left(\frac{\beta_1}{k}\right)(kx)$

$lin - lin$: $y = \beta_0 + \beta_1 x$ $\beta_1 = unit\ change\ in\ y\ for\ a\ 1\ unit\ change\ in\ x$

$lin - log$: $y = \beta_0 + \beta_1 \ln x$ $\frac{\beta_1}{100} = unit\ change\ in\ y\ for\ a\ 1\%\ change\ in\ x$

$log - lin$: $\ln y = \beta_0 + \beta_1 x$ $100\beta_1 = \%\ change\ in\ y\ for\ a\ 1\ unit\ change\ in\ x$

$log - log$: $\ln y = \beta_0 + \beta_1 \ln x$ $\beta_1 = \%\ change\ in\ y\ for\ a\ 1\%\ change\ in\ x$

$Partialling$: $For\ regression$: $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2$

$\hat{\beta}_1 = \frac{\sum_{i=1}^n \hat{r}_{i1} y_i}{\sum_{i=1}^n \hat{r}_{i1}^2}$, $where\ \hat{r}_{i1} = residual\ from\ OLS\ regression\ of\ x_1\ on\ x_2$

$These\ residuals\ are\ uncorrelated\ with\ x_2, and\ are\ estimates\ of\ x_1\ after\ effects\ of\ x_2\ have\ been$ $'partialled\ out.'\ \ Therefore, \hat{\beta}_1\ is\ simple\ relationship\ between\ y\ and\ x_1$ $after\ x_2 has\ been\ partialled\ out.$

Stat prop mult: Still linear in parameters, still random sampling, still zero condtional mean, still homoscedastic. Specific to multiple: no perfect multicollinearity in independent variables

OVB: $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2\ but\ we\ estimate\ y = \widetilde{\beta_0} + \widetilde{\beta_1} x_1\ and\ \ \widetilde{x_2} = \widetilde{\delta_0} + \widetilde{\delta_1} x_1\ then\ \widetilde{\beta_1} = \widehat{\beta_1} + \widehat{\beta_2}\widetilde{\delta_1}\ then\ E(\widetilde{\beta_1}) = \beta_1 + \beta_2\widetilde{\delta_1}$

$Including\ irrelevant\ vars\ maintains\ unbiasedness, but\ precision\ of\ estimates\ is\ bad.$

Multicollinearity: correlation among expl vars leads to bias in coefficients

$Var(\widehat{\beta_j}) = \frac{\sigma^2}{SST_j(1-R_j^2)}$ $where\ SST_j = \sum_{i=1}^n (x_{ij} - \overline{x_j})^2$ and

$R_j^2 = R^2\ of\ regression\ of\ x_j\ on\ all\ other\ expl\ vars$

$\hat{y} = \widehat{\beta_0} + \widehat{\beta_1} x_1 + \widehat{\beta_2} x_2$ and $\tilde{y} = \widetilde{\beta_0} + \widetilde{\beta_1} x_1$ $Var(\widehat{\beta_1}) = \frac{\sigma^2}{SST_1(1-R_1^2)}$ $Var(\widetilde{\beta_1}) = \frac{\sigma^2}{SST_1}$

$when\ \beta_2 \neq 0\ \ \widetilde{\beta_1}\ is\ biased\ \ \widehat{\beta_1}\ is\ unbiased$ and $Var(\widetilde{\beta_1}) \leq Var(\widehat{\beta_1})$

$when\ \beta_2 = 0\ \ \widetilde{\beta_1}, \widehat{\beta_1}\ both\ unbiased$ and $Var(\widetilde{\beta_1}) \leq Var(\widehat{\beta_1})$

Gauss-Markov Theorem: Given these five assumptions: 1 In population regression, y is linearly related to x and u. 2. Random sample of size n according to population model. 3. Sample outcomes of x are not all the same. 4. Expectation of error term is 0 given any value of x. 5. Error term has the same variance for any value of x. In a linear regression model, with expected errors of zero, uncorrelated errors, and equal variances, the Best Linear Unbiased Estimators are given by the OLS estimators.

Non-linear relationships in linear regressions: $add\ squared\ term$: $y = \beta_0 + \beta_1 x + \beta_2 x^2 + u$