Quiz 1. [10 points; Feb 22] Table 1 summarizes admissions to the graduate school at UC-Berkeley in 1973: it cross-classifies the admission decisions by type of department. Answer the following questions based on the table.

Table 1.

department type (X)	whether admitted (Y)	
	Yes	Not
A	971	547
B	591	1119
C	193	1105

Q1. [5/10] If 20 students admitted by type B departments were mistakenly recorded as "not admitted" in the table, how many people were in fact admitted by type B departments in total? How many people applied for type B depts?

Solution: (i) 591+20=611; (ii) 1710

Q2. [5/10] Write down the logistic regression model to study how admission decision at UC Berkeley (Y) is associated with type of department (X).

Solution: $Y = 1, 0$ for admitted, not; $X = 1, 2, 3$ for type A, B, C department. Let $\pi(x) = P(Y = 1|X = x)$. Logistic Regression Model:

$$\mathit{logit}\{\pi(x)\} = \alpha + \beta_x^X, \quad x = 1, 2, 3$$

- By Coding Scheme I. (SAS's default) $\beta_3^X \equiv 0$
- By Coding Scheme II. (R's default) $\beta_1^X \equiv 0$
- By Coding Scheme III. (ANOVA type) $\sum_{x=1}^{3} \beta_x^X \equiv 0$

Alternative modeling: $Y = 1, 0$ for admitted, not; $X_1 = 1, 0$ for type A, non-A department; $X_2 = 1, 0$ for type B, non-B department; Let $\pi(x_1, x_2) = P(Y = 1|x_1, x_2)$. Logistic Regression Model:

$$\mathit{logit}\{\pi(x_1, x_2)\} = \alpha + \beta_1 x_1 + \beta_2 x_2$$

- This is directly corresponding to Coding Scheme I. (SAS's default)
- The correspondence to Coding Scheme II. (R's default) is with $X_1 = 1, 0$ for type B, non-B department and $X_2 = 1, 0$ for type C, non-C department