

Rejecting Poolability in Simple Drug Stability Studies

Peng Liu
SAS Institute Inc.

Michael Crotty
SAS Institute Inc.

Liansheng Zhu
Novartis Pharmaceuticals

Abstract

While working on implementing a procedure to help practitioners determine the shelf life of a drug, we made some interesting findings. There are different interpretations about which approach to follow. We are aware of three approaches, which we compare using a simulation study. We will present the background of the study, descriptions of the three approaches, and comparisons among those approaches.

Key Words: stability analysis, hypothesis testing, sequential testing, simultaneous testing, model selection, federal guidelines

1. Introduction

A type of stability analysis is used to determine the expiration date of a drug under the long-term storage condition; see Q1E (2004). The typical data collected by the analysis include the measurements of certain drug attribute in capsules and the corresponding time that the measurements are taken. Linear regression models are used to fit the data. The analyst must select a model and extrapolate the fitted model until the predicted drug attribute crosses over some level, given a predetermined confidence level.

Unfortunately, there are different interpretations about what procedure is exactly needed to conduct the analysis. We will first present a set of possible models. Then according to different interpretations, we describe the corresponding procedures. All procedures are posed in reasonable and intuitive ways. We will not try to debate which procedure is the most appropriate, but we will try to illustrate the difference or the similarity in terms of the probability of rejecting poolability. The probability of rejecting poolability is directly related to the probability of determining a longer expiration.

2. Data and Models

It is common for the data to be collected in the following form:

Y	Batch	Time
Y_{11}	B_1	T_{11}
Y_{12}	B_1	T_{12}
...
Y_{21}	B_2	T_{21}
Y_{22}	B_2	T_{22}
...

The data table has three columns and multiple rows. Each row represents an observation, which includes the measurement of the drug attribute (denoted Y), the batch identification of the capsule, and the time that the measurement is taken.

All observations are independent. Following are four models, from which one will be selected and used for extrapolation.

The simplest model is

$$Y_{ij} = \beta_0 + \beta_1 T_{ij} + e_{ij}, \quad e_{ij} \stackrel{iid}{\sim} N(0, \sigma) \quad (1)$$

where subscript $i = 1, \dots, K$ indicates a batch out of total K , $j = 1, \dots, n_i$ indicates observations within each batch, and the e_{ij} 's are independently and identically distributed random normal errors with mean zero. We denote the distribution of normal errors by $N(0, \sigma)$, where σ is the standard deviation. This model implies that the way that the drug attribute varies with time is not affected by batches (i.e. poolability is not rejected).

The second model is

$$Y_{ij} = \beta_{0i} + \beta_1 T_{ij} + e_{ij}, \quad e_{ij} \stackrel{iid}{\sim} N(0, \sigma) \quad (2)$$

which only differs from Model (1) in that the intercepts of the regression lines are different for individual batches. This model implies that the drug attribute is different from batch to batch. However, the rate that the drug attribute varies with time is not affected by batches.

The third model is

$$Y_{ij} = \beta_{0i} + \beta_{1i} T_{ij} + e_{ij}, \quad e_{ij} \stackrel{iid}{\sim} N(0, \sigma) \quad (3)$$

which differs from Model (2) in that the slopes of the regression lines are different for individual batches as well. This model implies that the drug attribute and the rate that the attribute varies with time are both different from batch to batch.

The fourth model looks almost exactly the same as Model (3):

$$Y_{ij} = \beta_{0i} + \beta_{1i} T_{ij} + e_{ij}, \quad e_{ij} \stackrel{indep}{\sim} N(0, \sigma_i) \quad (4)$$

which only differs from Model (3) in that the random errors are distributed differently for individual batches. That is the distributions of random errors have different standard deviations for individual batches. We denote these by $N(0, \sigma_i)$, $i = 1, \dots, K$. This model implies that the variations among the drug attribute measurements, at a particular time point, are different from batch to batch.

3. Models and Testing Statistics

Although there are four models, only the inferences of two models are involved in model selection, regardless of which interpretation is adopted. The two models are (3) and (2). Table 1 and Table 2 are the symbolic representations of ANOVA tables of Models (3) and (2), respectively. Quantities of interest are indicated in the corresponding cells by a variable name. The labels of the sources follow the convention of SAS statistical procedures to represent Model (3) and Model (2). The meaning of these symbols and the formulae that compute them are available in many linear regression textbooks; Christensen (1996), Searle (1997), and Rawlings et al. (1998) are several among many others. Notice we put a superscript asterisk to indicate differences and relationships between some quantities.

The following important equalities exist between the models for Tables 1 and 2:

$$\begin{aligned} df_E^* &= df_C + df_E \\ SSE^* &= SSC + SSE. \end{aligned}$$

Table 1: ANOVA Table for Model (3)

Source	DF	Type I SS	Mean Square	F value	Pr > F
TIME					
BATCH	df_B	SSB	MSB	F_B	p_B
TIME*BATCH	df_C	SSC	MSC	F_C	p_C
ERROR	df_E	SSE	MSE		

Table 2: ANOVA Table for Model (2)

Source	DF	Type I SS	Mean Square	F value	Pr > F
TIME					
BATCH	df_B	SSB	MSB	F_B^*	p_B^*
ERROR	df_E^*	SSE^*	MSE^*		

4. Model Selection Procedures

We are aware of three procedures that interpret the description in Q1E (2004).

4.1 Procedure I

The first procedure (Procedure I) is described in the following steps:

1. If $p_B \geq 0.25$ and $p_C \geq 0.25$ in Table (1), accept the poolability of common slope and common intercept, select Model (1) and terminate this procedure. Otherwise, go to the next step.
2. If $p_B < 0.25$ and $p_C \geq 0.25$ in Table (1), accept the poolability of common slope, select Model (2) and terminate this procedure. Otherwise, go to the next step.
3. Reject the poolability of common slope and common intercept, and select Model (3).

This procedure was introduced in Ng and Tsong (1993). The intuition behind this procedure is a backward variable selection approach.

4.2 Procedure II

The second procedure (Procedure II) is described in the following steps:

1. If $p_C \geq 0.25$ in Table (1), accept the poolability of common slope and continue to the next step. Otherwise, select Model (3) and terminate this procedure.
2. If $p_B^* \geq 0.25$ in Table (2), accept the poolability of both common slope and common intercept, and select Model (1). Otherwise, reject the poolability of common intercept, and select Model (2).

The intuition behind the first step is to test the null hypothesis: $H_0 : \beta_{11} = \beta_{12} = \dots = \beta_{1K}$, for Model (3). And the intuition behind the second step is to test the null hypothesis: $H_0 : \beta_{01} = \beta_{02} = \dots = \beta_{0K}$, for Model (2).

4.3 Procedure III

The third procedure (Procedure III) is described in the following steps:

1. If $p_C \geq 0.25$ in Table (1), accept the poolability of common slope, and continue to the next step. Otherwise, select Model (3) and terminate this procedure.
2. If $Pr \left[F(df_B+df_C, df_E) > \frac{(SSB+SSC)/(df_B+df_C)}{SSE/df_E} \right] \geq 0.25$, where numbers are from Table (1), accept the poolability of both common slope and common intercept, and select Model (1). Otherwise, reject the poolability of common intercept, and select Model (2).

The intuition behind the first step is still to test the null hypothesis: $H_0 : \beta_{11} = \beta_{12} = \dots = \beta_{1K}$, for Model (3). But the intuition behind the second step is to test a simultaneous hypothesis: $H_0 : \beta_{11} = \beta_{12} = \dots = \beta_{1K}$ and $\beta_{01} = \beta_{02} = \dots = \beta_{0K}$, for Model (3).

5. Probability of Rejecting Poolability

The three procedures described in the previous section are closely related but different. Some analysts prefer to use one, while others may choose another. The three procedures all agree on how to reject the poolability of common slope. The difference lies in their probabilities of rejecting the poolability of both common slope and common intercept. Following are four test statistics, which define the following probabilities: p_{slope} , p_{I} , p_{II} and p_{III} :

$$\frac{SSC/df_C}{SSE/df_E} = F^{-1}(p_{\text{slope}}, df_C, df_E) \tag{5}$$

$$\frac{SSB/df_B}{SSE/df_E} = F^{-1}(p_{\text{I}}, df_B, df_E) \tag{6}$$

$$\frac{SSB/df_B}{SSE^*/df_E^*} = F^{-1}(p_{\text{II}}, df_B, df_C + df_E) \tag{7}$$

$$\frac{(SSB + SSC)/(df_B + df_C)}{SSE/df_E} = F^{-1}(p_{\text{III}}, df_B + df_C, df_E), \tag{8}$$

where $F^{-1}(p, df_1, df_2)$ is the p^{th} quantile of the F distribution with degrees of freedom df_1 and df_2 . Notice SSB , SSC , and SSE are all independent quantities if assumptions of linearity and normality hold, so they can be any positive numbers. Meanwhile, $p_{\text{slope}} = 1 - p_C$, $p_{\text{I}} = 1 - p_B$, and $p_{\text{II}} = 1 - p_B^*$. They are probabilities, rather than p-values, so the corresponding poolability by following the particular procedure will be rejected if the probability value is greater than or equal to 0.75. All procedures will reject the common slope if $p_{\text{slope}} \geq 0.75$. Procedure I will reject both poolabilities of slope and intercept if $p_{\text{I}} \geq 0.75$. Procedure II will reject both poolabilities of slope and intercept if $p_{\text{II}} \geq 0.75$. Procedure III will reject both poolabilities of slope and intercept if $p_{\text{III}} \geq 0.75$. Also notice $df_B = df_C$, because both equal $K - 1$. Furthermore, because we only need to compare p_{I} , p_{II} , and p_{III} , when poolability of common slope is confirmed, we can assume

$$p_{\text{slope}} < 0.75.$$

After some algebra, we can rewrite the left hand side of Equation (7) in terms of Equation (5) and Equation (6), and apply $df_B = df_C$. The result is:

$$\frac{F^{-1}(p_I, df_B, df_E) \times (df_B + df_E)}{F^{-1}(p_{\text{slope}}, df_B, df_E) \times df_B + df_E} = F^{-1}(p_{II}, df_B, df_B + df_E).$$

Therefore,

$$p_{II} = F \left(\frac{F^{-1}(p_I, df_B, df_E) \times (df_B + df_E)}{F^{-1}(p_{\text{slope}}, df_B, df_E) \times df_B + df_E}, df_B, df_B + df_E \right),$$

where $F(x, df_1, df_2)$ is the distribution function of the F distribution with degrees of freedom df_1 and df_2 . Now rewrite the left hand side of Equation (8) in terms of Equation (5) and Equation (6), and apply $df_B = df_C$. The result is:

$$\frac{1}{2} \left[F^{-1}(p_{\text{slope}}, df_B, df_E) + F^{-1}(p_I, df_B, df_E) \right] = F^{-1}(p_{III}, 2df_B, df_E).$$

Therefore,

$$p_{III} = F \left(\frac{1}{2} \left[F^{-1}(p_{\text{slope}}, df_B, df_E) + F^{-1}(p_I, df_B, df_E) \right], 2df_B, df_E \right).$$

Hence, we have expressed both p_{II} and p_{III} in terms of p_{slope} , p_I , df_B , and df_E . Therefore, we can compare the probability of rejecting poolability, regardless what the true model is, given those four quantities. The next section conducts a simulation study.

6. Simulation Study and Discussions

The design of the simulation involves only the ranges and variations of four quantities, which are p_{slope} , p_I , df_B , and df_E . As we have shown in the previous section, the true model and the true distribution are irrelevant in comparing p_{II} and p_{III} , given p_{slope} , p_I , df_B , and df_E . In this simulation, we assume p_{slope} is from a random Uniform(0, 0.75), p_I is from a random Uniform(0, 1), df_B is from a random Uniform(2, 12), and df_E is from a random Uniform(1, 20). Hence, we guarantee that there are at least three batches; the poolability of the common slope cannot be rejected; there is at least one degree of freedom for errors. The simulation has 10000 samples. The comparisons of rejecting poolability of a common slope and a common intercept are tabulated in the following three tables.

	Reject by Procedure I	
Reject by Procedure II	No	Yes
No	6458	44
Yes	1113	2385

	Reject by Procedure I	
Reject by Procedure III	No	Yes
No	7568	1149
Yes	3	1280

	Reject by Procedure II	
Reject by Procedure III	No	Yes
No	6476	2241
Yes	26	1257

It is interesting to see that, in this simulation, the probability of rejecting poolability by following Procedure II is greater than that by following Procedure I. Meanwhile, Procedure III mostly agrees with the decision of not rejecting poolability from Procedure I. In other words, if the poolability of common slope is accepted by Procedure I, then Procedure I and III are very much likely to agree on the decision on pooling both common slope and intercept. However, by nearly half of the chance, Procedure III does not agree with decision of rejecting poolability from Procedure I.

The consequence of accepting poolability is related to the likelihood of a procedure determining a longer expiration. In general, the more likely that poolability is accepted, the longer the determined expiration will be. When we are talking about determining the expiration date, we would like to point out the estimation procedure introduced by Ng and Tsong (1993) uses Model (4). In other words, if both poolabilities of slope and intercept are rejected, the SAS STAB macro uses Model (4) instead of Model (3) which one might expect. This choice may largely shorten estimated expiration dates, when poolability is rejected. It is also interesting to think about the real world scenarios, which is that three procedures are not as different as they appear in the above simulation, because the distributional assumptions on p_{slope} , p_{I} , df_B , and df_E may not be realistic.

References

- Center for Drug Evaluation and Research (U.S.) & Center for Biologics Evaluation and Research (U.S.) & International Conference on Harmonisation, (2004), *Guidance for Industry : Q1E Evaluation of Stability Data*, U.S. Dept. of Health and Human Services, Food and Drug Administration, Center for Drug Evaluation and Research: Center for Biologics Evaluation and Research, Rockville, MD.
- Christensen, R., (1996), *Plane Answers to Complex Questions: The Theory of Linear Models*, Springer.
- Searle, S. R., (1997), *Linear Models*, Wiley.
- Rawlings, J. O., Pantula, S. G., and Dickey, D., (1998), *Applied Regression Analysis: A Research Tool*, Springer.
- Ng, M., and Tsong, Y., (1993), *A PC/SAS Stability Analysis Software*, Proceedings of the SAS Users Group International Conference, Volume 18, pp. 1141-1148.