Auteursrechterlijke overeenkomst

Opdat de Universiteit Hasselt uw eindverhandeling wereldwijd kan reproduceren, vertalen en distribueren is uw akkoord voor deze overeenkomst noodzakelijk. Gelieve de tijd te nemen om deze overeenkomst door te nemen, de gevraagde informatie in te vullen (en de overeenkomst te ondertekenen en af te geven).

Ik/wij verlenen het wereldwijde auteursrecht voor de ingediende eindverhandeling met

Titel: Vlaamse normering van de Purdue Pegboard voor leerlingen binnen OV2 (BuSO)
Richting: master in Applied Statistics Jaar: 2008

in alle mogelijke mediaformaten, - bestaande en in de toekomst te ontwikkelen - , aan de Universiteit Hasselt.

Niet tegenstaand deze toekenning van het auteursrecht aan de Universiteit Hasselt behoud ik als auteur het recht om de eindverhandeling, - in zijn geheel of gedeeltelijk -, vrij te reproduceren, (her)publiceren of distribueren zonder de toelating te moeten verkrijgen van de Universiteit Hasselt.

Ik bevestig dat de eindverhandeling mijn origineel werk is, en dat ik het recht heb om de rechten te verlenen die in deze overeenkomst worden beschreven. Ik verklaar tevens dat de eindverhandeling, naar mijn weten, het auteursrecht van anderen niet overtreedt.

Ik verklaar tevens dat ik voor het materiaal in de eindverhandeling dat beschermd wordt door het auteursrecht, de nodige toelatingen heb verkregen zodat ik deze ook aan de Universiteit Hasselt kan overdragen en dat dit duidelijk in de tekst en inhoud van de eindverhandeling werd genotificeerd.

Universiteit Hasselt zal mij als auteur(s) van de eindverhandeling identificeren en zal geen wijzigingen aanbrengen aan de eindverhandeling, uitgezonderd deze toegelaten door deze overeenkomst.

Ik ga akkoord,

DUAN, Yuanxing

Datum: 5.11.2008

Vlaamse normering van de Purdue Pegboard voor leerlingen binnen OV2 (BuSO)

Yuanxing Duan

promotor :
Prof. dr. Herbert THIJS,
Dr. Raf MEESEN,
Drs. Tetyana KADANKOVA

Universiteit Maastricht

universiteit hasselt

Eindverhandeling voorgedragen tot het bekomen van de
graad master in Applied Statistics

Content

0. Abstract

The aim of this study is to test the effect of age, gender, dominant hand of students in province Limburg Belgium on Purdue Pegboard Test which was developed to test employees with requirement on fine and gross motor dexterity and coordination. Linear regression, multivariate regression and general linear mixed model were approached, the results showed that the baseline of the profiles is influenced by age, gender, but dominant hand. As 4 different subtests were set, several models were fitted.

Key words and phrases: Purdue Pegboard Test, linear regression, multivariate regression, general linear mixed model.

1. Introduction

The Purdue Pegboard Test was first developed by Joseph Tiffin, Ph.D., an Industrial Psychologist at Purdue University in 1948. Since that time, this device has been used extensively to aid in the selection of employees for jobs that require fine and gross motor dexterity and coordination. It measures gross movements of hands, fingers and arms, and fingertip dexterity as necessary in assembly tasks. The device has 4 cups across the top, 2 vertical rows of 30 small holes down the center.

Purdue Pegboard Test tests dexterity. It measures two types of activities: gross movements of hands, fingers and arms, and "fingertip" dexterity in an assembly task which involves sequential insertion of pegs and assembly of pegs, collars and washers.

In this experiment, a total of 321 people (134 females and 187 males) aged from 12 to 24 were asked to do four subtests, each should be done as many as possible 3 times. 4 subtests are dominant hand in 30 seconds, non-dominant hand in 30 seconds, bimanual in 30 seconds and assembly in 60 seconds. There are 33 missing values in the dataset, because some of the subjects only did one or two subtests.

Because the response of interest number which subjects put is count number, based on the rule of thumb for count data, if the response of interest is larger than 7, a linear model which assuming normal distribution instead of Poisson distribution will be approached (Agresti, 2002). A linear regression with the mean of outcomes under each subtest was fitted as the first attempt to see the effect of each covariate, it may not be a good approximation because the response of interest is a count number, and the mean of count doesn't have any meaning. So a more advanced multivariate regression model will be fitted, because the dataset is equally spaced, and the variable 'time' indicates the order of which the subject did the experiment. Multivariate regression is efficient at balanced data. In order to achieve more accurate results, a general linear mixed model, which is more advanced than multivariate regression will be fitted.

2. Methodology

2.1 Variable description

Table 1 Overview of variables in the dataset.

Variable	Description	Nature of the variable
Subject	Number of subjects	Identity
Response	Number that each subject makes	Continuous
Age	Age of subjects	Continuous
Sex	Gender of subjects	Categorical (0,1)
Time	Time variable	Ordered categorical
Hand	Dominant hand	Categorical (0,1)
Subtest	Different subtests	Categorical (1 to 4)

There are 321 subjects aged from 12 to 24 in the whole experiment, who were asked to do each subtest for 3 times, the response of interest is the number that each subject made in the experiment. The covariate time is the time order that a certain subject did the experiment, it is an ordered categorical covariate. The covariate sex was set to be dummy variable with 134 females denoted by 0 and the other 187 males were denoted by 1. Also the covariate hand denoted the dominant hand of subjects, where 1 represented right hand, and with 0 presented left hand.

2.2 Exploratory data analysis

Because data exploration is extremely helpful as additional tool in the selection of appropriate models, exploratory data analysis is approached. To explore the mean structure, variance structure and correlation structure, graphical techniques were explored. The mean evolution shows the profile of the mean over time. For balanced data, averages can be calculated for each occasion separately. The variance function equals $\sigma^2(t) = E\left[Y(t) - \mu(t)\right]^2$, hence, an estimate for $\sigma^2(t)$ can be obtained from applying any of the techniques described for exploring the mean structure to squared residuals r_{ij}^2.

2.3 Statistical models

2.3.1 Linear regression model

A regression model is a formal means of expressing the two essential ingredients of a statistical relation: a tendency of the response variable Y to vary with the predictor variable X in a systematic fashion, and a scattering of points around the curve of statistical relationship (Kutner et al, 2005).

A single model with one covariate is used as an example.

$$Y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$$

Where Y_i is the value of the response for the ith subject, β_0 and β_1 are parameters, x_i is a known constant value for subject i, ε_i is a random error term with mean 0 and variance σ^2. It assumes that Y_i follows a normal distribution with mean $\beta_0 + \beta_1 x_i$ and variance σ^2.

In this study, linear regression could only be fitted separately by each subtest, with mean as response, because it assumes the independence between subjects.

2.3.2 Multivariate regression model

The multivariate regression model is equivalent to the classical regression model except that it assumes independence between subjects and allows for within subjects correlation of the repeated measurements.

The multivariate regression assumes a completely balanced data, equal number of measurement per subjects taken at fixed time points. This is in contrast to the dataset in which there are unequal numbers of measurement per patient. In this study maximum likelihood would be used in order to make the reduction test simpler.

A multivariate regression model was fitted separately, so four multivariate regression models were fitted. In subtest 1, let Yi be the vector of 3 repeated measurements for the ith subject: $Y_i = (Y_{i1} \quad Y_{i2} \quad Y_{i3})'$. The general multivariate model assumes that Yi satisfies a regression model:

$Y_i = X_i\beta + \varepsilon_i$ with

- X_i: Matrix of covariates
- β: Vector of regression parameters
- ε_i: Vector of error components, $\varepsilon_i \sim N(0, \Sigma)$

We then have the following distribution for Y_i: $Y_i \sim N(X_i\beta, \Sigma)$, the mean structure is modeled $X_i\beta$ as in classical linear regression models. Usually, Σ is just a general $(n \times n)$ covariance matrix, however special structure for Σ like first-order autoregressive and compound symmetric structures can be assumed.

Assuming independence across individuals, β and the parameters in Σ can be estimated by maximizing

$$L_{ML} = \prod_{i=1}^{N} \left\{ (2\pi)^{-n/2} |\Sigma|^{-1/2} \exp\left(-\frac{1}{2} (y_i - X_i \beta)' \Sigma^{-1} (y_i - X_i \beta) \right) \right\}$$

Inference is based on classical maximum Likelihood ratio test.

2.3.3 General linear mixed model

The General Linear Mixed Model is expressed by the following formal equation:

$$Y_i = X_i\beta + Z_i b_i + \varepsilon_i$$

with $b_i \sim N(0,D)$; $\varepsilon_i \sim N(0,\sigma^2 I_{ni})$; $b_1,\ldots b_N$ and $\varepsilon_1,\ldots,\varepsilon_N$ independent.

X_i and Z_i are the design matrices and the vectors β and b_i represent regression parameters: β contains all parameters which are the same for all clusters (fixed effects) and b_i contains the cluster-specific parameters (random effects). Measurement error components are represented by the vector ε_i. It is a stretch to the linear regression with the random effects, so that the variability could be well measured.

2.4 Software package

All statistical analyses were performed in SAS 9.1 (SAS Institute Inc.). Significance level of all tests was set on $\alpha = 0.05$.

3. Statistical analysis

3.1 Exploratory Data Analysis

3.1.1 Individual profiles

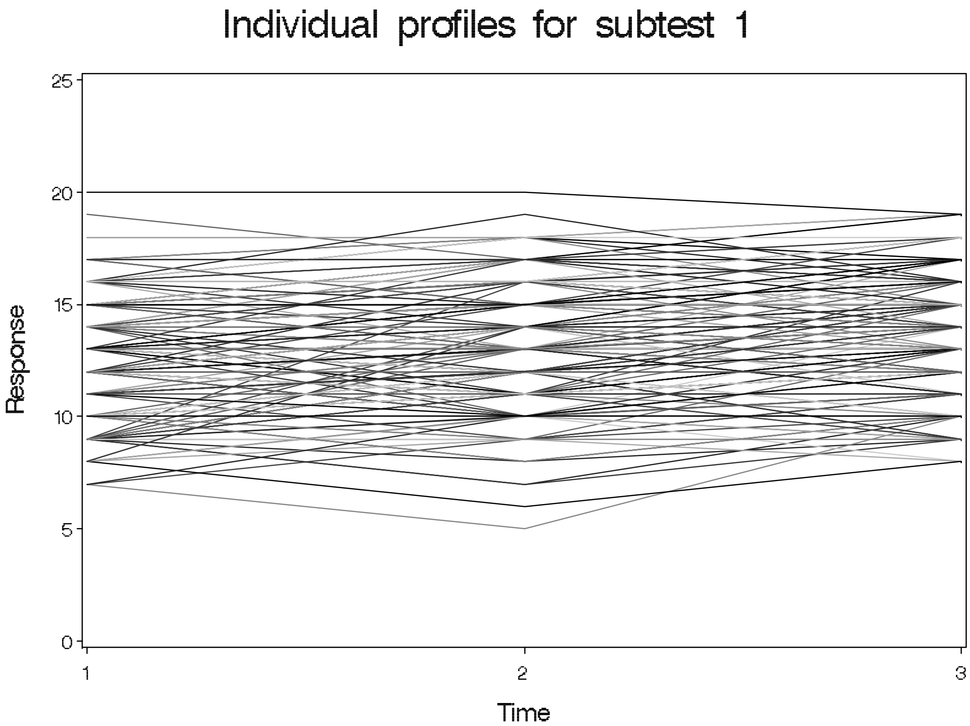


Figure 3 Individual profiles of subtest 1

From figure 1 and figure A1 in appendix, it showed that each subject did different results on the different time point. Great within variability showed in all the 4 subtests, especially in the fourth subtest, the smallest value is 2 and the largest value is 51, and most values are between 10 and 40. With this balanced dataset, each subject has 3 measurements in each subtest, because the measurement value at each time point is different, subjects do not have constant line in the whole process of each subtest.

3.1.2 Exploring the mean structure

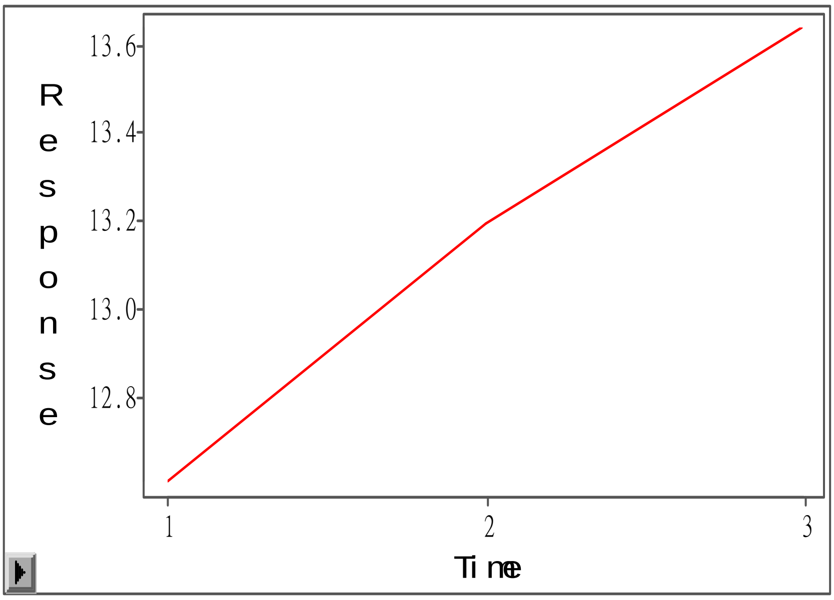


Figure 2 the mean structure of response for subtest 1

Figure 2 and figure A2 in appendix clearly showed that the means of the 4 subtests increased as the time increased. And because in the first 3 subtests subjects were asked to do the experiment in 30 seconds and in the fourth, subjects were given 1 minute to finish it, the mean in the fourth subtest is much larger than the other three.

3.1.3 Exploring the variance structure

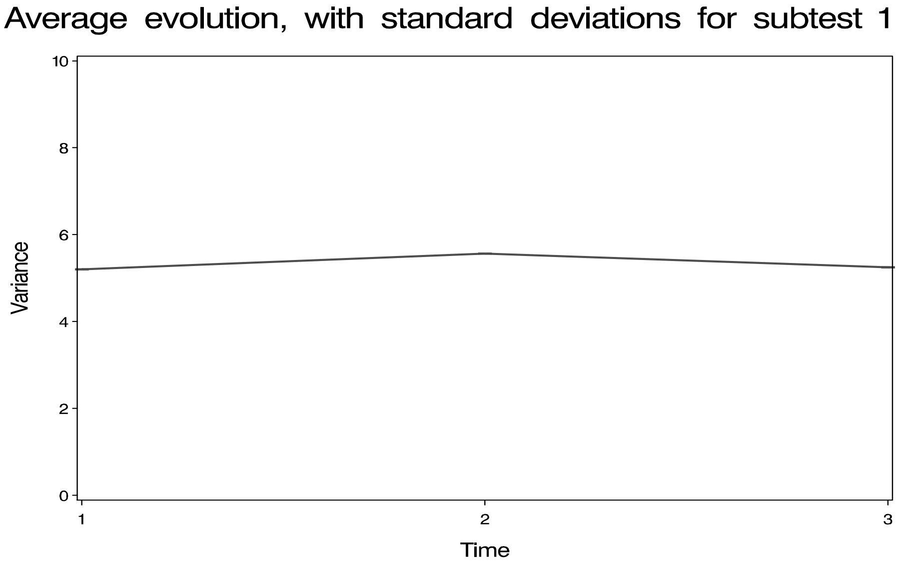


Figure 3 the variance structure of subtest 1

From figure 3 and figure A3 in appendix, it showed that there is no much variability in the first 3 subtests, the variance is around 6, but for subtest 4, much variability was found, the variance at time point 1 is over 50, at time point 3, it increased to over 70, it indicates that there is larger variability in subtest 4.

3.1.4 Exploring the correlation structure

Table 2 Correlation structure of subtest 1

Pearson Correlation Coefficients, N = 321			
	time1	time2	time3
time1	1.00	0.78	0.77
time2		1.00	0.79
time3			1.00

From table 2 and table A1, it shows that the correlation between each two time points were around 0.8, it interpreted that time maybe played little effect on the response because subjects did experiments in close time difference. In subtest 1 and 4, the correlation between time point 1 and time 2 is a little bit larger than that between time point 2 and 3, but in subtest 2 and 3, the correlation between time point 1 and time 2 is a little bit smaller than that between time point 2 and 3. As all the correlations in the table are all around 0.8, a compound symmetry covariance structure may be appropriate.

3.2 Result

3.2.1 Linear regression

Table 3 Results from linear regression

	Subtest 1		Subtest 2		Subtest 3		Subtest 4	
Effect	P.E.	Pr > |t|	P.E.	Pr > |t|	P.E.	Pr > |t|	P.E.	Pr > |t|
Intercept	10.75	<.0001	10.00	<.0001	7.12	<.0001	15.12	<.0001
Age	0.16	0.0015	0.13	0.0066	0.13	0.0256	0.57	0.0011
Sex	-0.54	0.0419			-1.47	<.0001		

From table 3, it can be concluded that in all the 4 subtests, age has certain effect, because the p-values are smaller than 0.05, the significance level. With positive parameter estimates, it indicated that age has a positive to the response of interest. The covariate sex is significant in subtest 1 and 3, it means that males would make less than females. The Shapiro-Wilk tests were approached in the 4 subtests, the p-values were all around 0.88, which indicated that there is no enough evidence to reject the null-hypothesis, so the normality held.

Because the response of interest in linear regression models is the mean, which is calculated by ignoring the effect of time, a more advanced model with considered the correlation within time, multivariate regression model, would be approached in the following.

3.2.2 Multivariate regression

In subtest 1, a saturated multivariate regression model with all the main effects and all the possible interactions used as covariates was obtained. Then a similar model with one less interaction, which provided the largest p-value in the saturated model, was fitted, using the likelihood ratio test which was mentioned before in the methodology part, the mean structure of the model was reduced without losing the efficiency into a model with one less interaction, as it follows the Chi-square distribution. P-value which is larger than 0.05 (the significance level) indicates that there is no significant different between the two models. Within the same method step by step, a model with only the main effects was proved that it has the same efficiency as the saturated model. Using the LR test again, a reduced mean effect model with only covariate age and time was obtained, a step further test was approached, but the value of -2*log-likelihood is extremely large, in the contrast, the degrees of freedom are rather small. So the further model selection is about the covariance structure, because for the mean structure, it is the simplest and within the same efficiency among all models were fitted. As the unstructured covariance structure is the most complicated, likelihood ratio test could also be applied to select the covariance structure, as the reduce main effect model was under unstructured covariance structure and our dataset is balanced, a first-order autoregressive (AR(1)) and Compound Symmetry were applied, under likelihood ratio test Compound Symmetry with p-value equaled 0.7358 was chosen, AR(1) with p-value much smaller than 0.05 was eliminated. The corresponding models with covariance structure, number of parameters and log-likelihood were listed in table 4.

Table 4 model reduction for subtest 1

Model	Covariance structure	No. Param	-2 log likelihood
Saturated	Unstructured	38	3622.8
Main effect	Unstructured	14	3635.2
Reduced main effect	Unstructured	10	3637.9
Covariance reduced 1	1st-order Autoregressive	6	3693.0
Covariance reduced 2	Compound Symmetry	6	3639.9

The SAS output of the multivariate regression for subtest 1 is shown in table 5 below. Because the computational limitation of testing the multivariate normality, three univariate tests were done one by one, the p-values of Shapiro-Wilk test are 0.59, 0.59 and 0.18 for the each of the response separately. Age is significant with p-value equals

0.001, as the parameter estimate of age is 0.1486, it is positive which indicated that the elder the subjects are, the more they can make in the experiment, the same result as in linear regression, but this conclusion should be based on the certain age interval in this experiment, which is from 12 to 24. As the parameter estimates of time point 1 and time point 2 are negative, compared with time point 3 which is the base line, subjects made less at time point 1 and 2, it interpreted that as the time went by, subjects could make more.

Table 5 SAS output for subtest 1

Effect	time	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept		11.1251	0.7636	322	14.57	<.0001
age		0.1486	0.04466	319	3.33	0.0010
time	1	-1.0313	0.08594	638	-12.00	<.0001
time	2	-0.4357	0.08594	638	-5.07	<.0001
time	3	0	.	.	.	.

Within the same method as subtest 1, SAS output for the other three subtests are listed in table A2 in appendix. Almost the same results as subtest 1 were obtained, the elder the subjects are, the more they can make, but in subtest 4, the parameter estimate of age is 0.5227 which is much larger than 0.14 in subtest 1, 0.12 in subtest 2 and 0.18 in subtest 3. Because in subtest 4, subjects were given 1 minute to finish the experiment and they can put any one of the three objects into the hole on the board as they like.

The parameter estimates of time in all the four subtests owned the same property that the parameter estimates at time point 1 and 2 are negative, as interpreted before, it means that as the time went by, subjects can make more. The only difference from these 4 subtests is in subtest 3 the p-value of covariate sex is smaller than 0.05 which indicates that sex can also affect the number that subjects made, as males were set to be the base line which is different from the linear regression part, and the parameter estimate is 1.19 which is positive, so in the subtest 3 when subjects were asked to do the experiment with two hands simultaneously, females can make 1.19 more on average than males.

For subtest 2, the normality of the first time point held, with p-value equals 0.24, but the other two rejected the null-hypothesis of normality with p-values equal 0.02 and 0.0002. And in subtest 3, normality held at time point 1 and 3, but time point 2 with p-value equals 0.002. In subtest 4, all the three rejected the normality null-hypotheses which had p-value less than 0.05.

A model with all the four subtests together was fitted, it is more complicated than the separate part although there is no much difference in the mean structure, the covariance structure is different that’s because the variance and covariance matrix is different, in the separate part, each subject have 3 measurement, but here each subject has 12 measurement, the variance and covariance matrix changed from 3 by 3 to 12 by 12 because of balanced design of the experiment. In the model selection part, the same likelihood ratio test as in the separate part was approached. The final model contained the main effect of age, sex and time, the unstructured covariance structure was applied as in the covariance reduction part, and the first-autoregressive and compound symmetry provide highly significant p-value. The SAS output for the final model was shown below.

Table 4 SAS output for the combined model

Effect	sex	time	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept			8.4004	0.6070	320	13.84	<.0001
sex	0		0.4792	0.1852	319	2.59	0.0101
sex	1		0	.	.	.	.
age			0.1007	0.03491	318	2.88	0.0042
time		1	-0.3718	0.04867	319	-7.64	<.0001
time		2	-0.04395	0.04501	319	-0.98	0.3296
time		3	0	.	.	.	.

The result is similar as the third subtest, parameter estimates of time at time point 1 and 2 are negative, so the same result as in the separate part is obtained. And the same result of sex is obtained, female did better than male. The parameter estimate of age is still positive, also the same result as before. Within the Shapiro-Wilk test univariately, some held normality, in the contrast, some rejected normality.

3.2.3 General linear mixed model

Within the general linear mixed model, in the random-effect selection part, the model without random effects, with random intercept and with random intercept and slope shared the same -2log-likelihood value. Therefore a general linear mixed model which is more advanced than multivariate regression and would be a more correct approach, could not be used because of computational limitation. For this the second most advanced model was chosen as the final one, and the interpretations were all based on the multivariate regression model taking into consideration.

4. Conclusion and discussion

The experiment is about putting the pegs, collars and washers into the holes on the Pegboard which was invented by Purdue University in 1948, which was used to aid in the selection of employees for jobs that require fine and gross motor dexterity and coordination. The main purpose of this study is to tell the difference of age, sex and dominant hand of the certain subject in the whole process. Because there are 4 subtests, which are dominant hand, non-dominant hand, bimanual and assembly, in the first three, subjects were given 30 seconds to put pegs, but in the fourth subjects were given a whole minute to put any of the three into the holes.

In linear regression, which used the mean of three measurements as the response, age does have effect on the outcome and the parameter estimates are all positive. In subtest 1 and 3, females did better than males, but there is no difference between gender in subtest 2 and 4. And normality held for all the four subtests.

Multivariate regression models were approached for the statistical analysis, because the dataset was balanced designed. Each subject was measured 3 times in each subtest, so that each subject had 12 measurements in total. 5 models were fitted, four were for the four subtests separately, and one was for the dataset which combined the four subtests. Likelihood ratio test was applied in the model selection part to reduce the mean and covariance structure. Although only in subtest 1, all the normality held, because there are 3852 observations in total, it is large enough, as for small sample, the normality always holds, but for large sample, the normality is always been denied, nevertheless, in such large sample, little attention should be paid on it.

In all the 5 models, there seemed no interaction effect between any two of the main effects. In subtest 1, 2 and 4, only age had effect on the number of measurement, in subtest 3, sex also affected the response of interest which is different the results of linear regression, and the results showed that females did better in the third subtest. In model with all the 12 measurements, age and sex both had effect on the response of interest. As the scientists suggested that females with long and slim fingers always did better than males, within the statistical analysis, the same result was obtained, totally females did better than males. And age affected the response of interests in all the five models, it is the same result as linear regression, but this conclusion should be made in the certain age interval in the experiment which is from 12 to 24 in the dataset. In all the five models, whether the dominant hand of subjects is left or right did not affect the response of interest.

Because statistically, general linear mixed model could not be used because of computational limitation, no conclusion was based on it. There may be a statistical problem on that, as general linear mixed model has sophisticated calculation, it is difficult to find the exact problem.

As for the count response of interest, the first distribution assumption should be Poisson, but in this study in each subtest subjects were asked to do the experiment three times, a Generalized Linear Mixed Model with log-like could be fitted, but because of the computational limitation, a model with both random intercept and random slope could not be fitted due to the negative definition of G matrix.

In summary, based on the results of multivariate regression, the elder the subjects are, the more pegs they can put into the holes on the Pegboard, but it should be concluded under the certain age interval in Limburg, Belgium. And females did better than males, it is the same as the scientists advised. For the effect of dominant hand, none of the five models showed it, so there is no effect of the dominant hand of subjects.

5. Reference

1. http://www.wisdomking.com/product158206c75034.html revised on 15 April, 2008

2. Verbeke, G., and Molenberghs, G. (2000). *Linear Mixed Models for Longitudinal Data.* New York: Springer-Verlag.

3. Verbeke, G., and Molenberghs, G. (2007). *Project: Longitudinal Data Analysis.* Diepenbeek: Universiteit Hasselt, Unpublished Course notes.

4. Johnson, R. and Wichern, D. (2007). *Applied Multivariate Statistical Analysis. (6th ed.).* NJ: Prentice Hall.

5. Molenberghs, G. (2007). *Multivariate Data Analysis.* Diepenbeek: Universiteit Hasselt, Unpublished Course notes.

6. Kutner, M., Nachtscheim, C., Neter, J. And Li, W., (2005), *Applied Linear Statistical Models (5th ed.)* McGraw-Hill/Irwin

7. Agresti, A. (2002). *Categorical Data Analysis (2nd Edition).* NY: Wiley InterScience.

Appendix

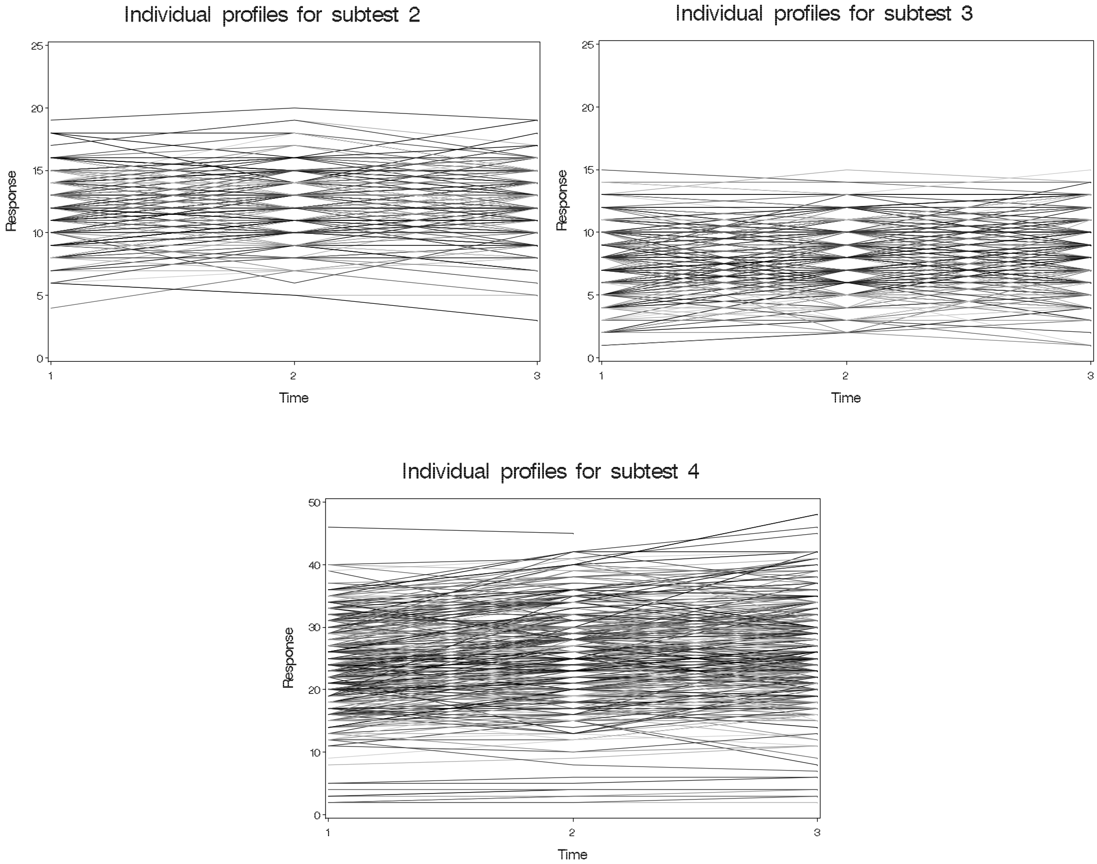
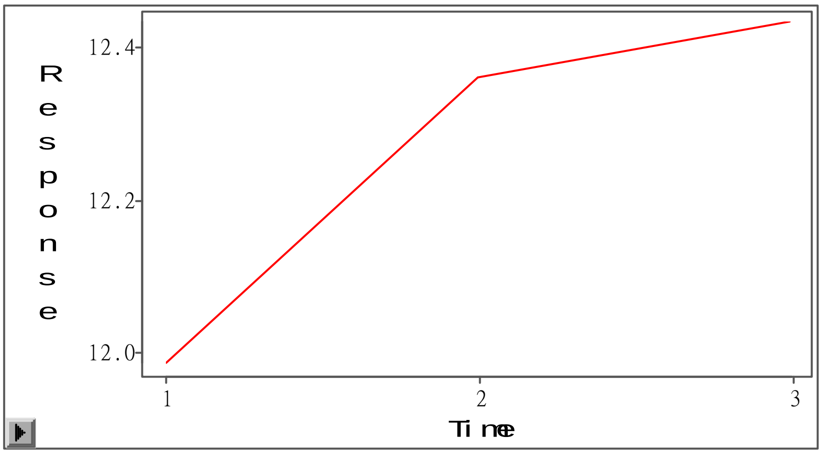
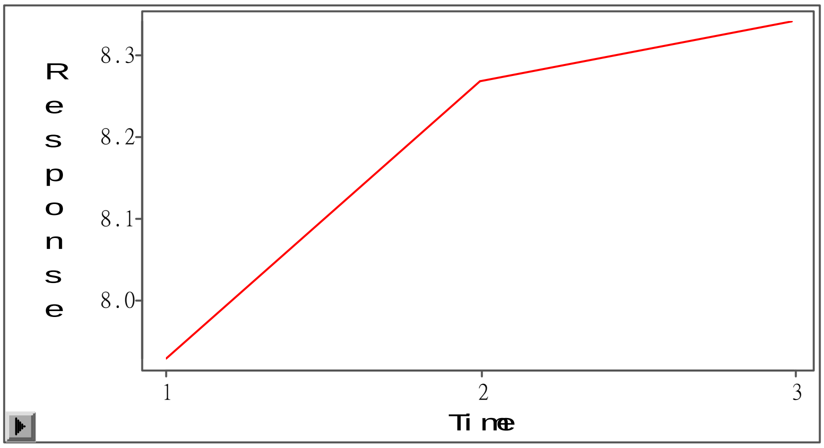


Figure A1 Individual profiles in subtest 2, 3 and 4





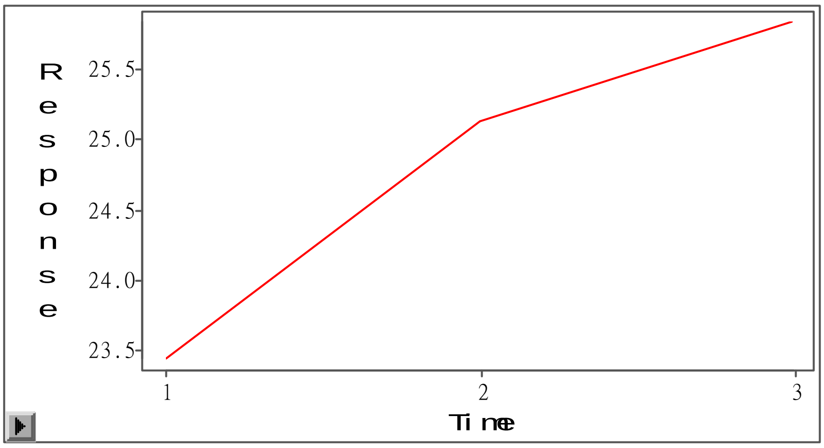
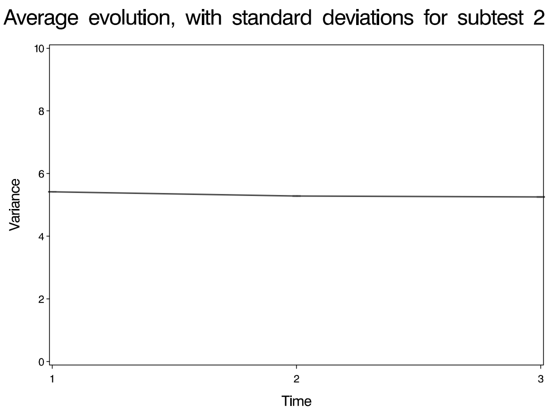
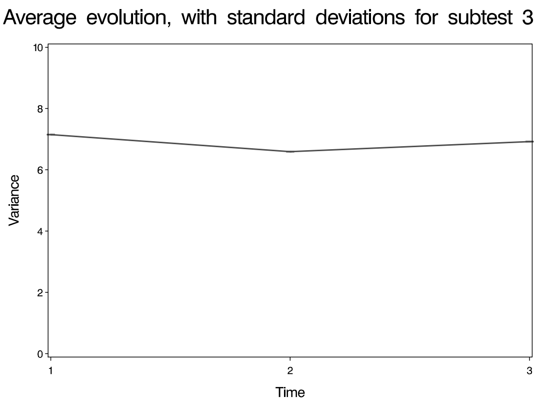


Figure A2 Mean structure of subtest 2, 3 and 4





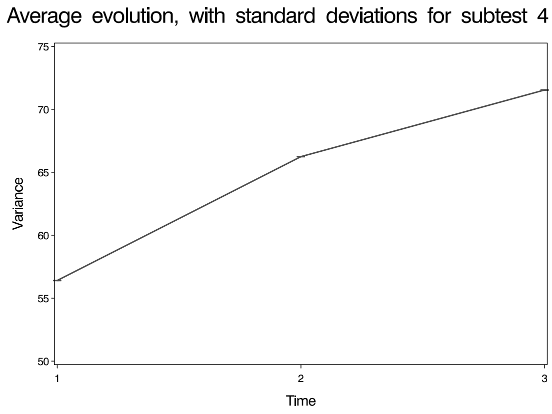


Figure A3 Variance structure for subtest 2, 3 and 4

Table A1 Correlation structure for subtest 2, 3, 4

Subtest 2

Pearson Correlation Coefficients, N = 320			
	time1	time2	time3
time1	1.00	0.79	0.80
time2		1.00	0.82
time3			1.00

Subtest 3

Pearson Correlation Coefficients, N = 320			
	time1	time2	time3
time1	1.00	0.78	0.82
time2		1.00	0.79
time3			1.00

Subtest 4

Pearson Correlation Coefficients, N = 320			
	time1	time2	time3
time1	1.00	0.89	0.88
time2		1.00	0.91
time3			1.00

Table A2 SAS output for subtests 2, 3 and 4 in multivariate regression

Subtest 2

Effect	time	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept		10.3505	0.7711	320	13.42	<.0001
age		0.1222	0.04510	318	2.71	0.0071
time	1	-0.4434	0.08144	636	-5.44	<.0001
time	2	-0.06289	0.08144	636	-0.77	0.4403
time	3	0	.	.	.	.

Subtest 3

Effect	sex	time	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept			4.7475	0.8670	320	5.48	<.0001
time		1	-0.4182	0.09376	636	-4.46	<.0001
time		2	-0.07233	0.09376	636	-0.77	0.4407
time		3	0	.	.	.	.
sex	0		1.1914	0.2641	318	4.51	<.0001
sex	1		0	.	.	.	.

Effect	sex	time	Estimate	Standard Error	DF	t Value	Pr > |t|
age			0.1825	0.04979	318	3.67	0.0003

Subtest 4

Effect	time	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept		17.0060	2.7025	320	6.29	<.0001
age		0.5227	0.1578	316	3.31	0.0010
time	1	-2.3994	0.2237	317	-10.73	<.0001
time	2	-0.7138	0.1957	317	-3.65	0.0003
time	3	0	.	.	.	.