EVALUATION AND COMPARISON OF METHODS OF MEASUREMENTS

DAY 3 - 1

Laboratory methods

Niels Trolle Andersen and Mogens Erlandsen

mogens@biostat.au.dk

Department of Biostatistics

TODAY

PART 1

Assay methods

- **Measurement error models**
- **Laboratory methods - Calibration**
- **Comparing assay methods**

EXERCISES

PART 2

- **KAPPA**
- **DESIGN CONSIDERATIONS**
- **NO EXERCISES**
- **EVALUATION**

Measurement error model

Notation: (i'th individual)

X_i true value

x_i observed value

u_i random error

Classical: $x_i = X_i + u_i$, X_i and u_i independent

Observed variable (x) measured with error

Berkson: $X_i = x_i + u_i$, x_i and u_i independent

Group of individuals assigned the same value (x), but true value (X) vary among individuals (compare with residual confounding, e.g. smoking)

Laboratory methods

Laboratory methods:

•Direct measurement (e.g. weight)
- rarely used (often slow and expensive)

•Indirect measurement – requires calibration
- fast, cheap

Assay methods

Some sources to measurement errors:

•Matrix

•Calibration

•Treatment (of test and calibration material)

Assay methods

Calibration:

Indirect method (e.g. HPLC, RIA, etc.) to measure y

Task: Calculate the corresponding value of X (e.g. a concentration)

Basic assumption: Existence of a measurement function f such that $y = f(X)$

Problem: The function f is (in general) unknown

”Repeatablility” – within assay

”Reproducability” – between assay (recalibration)

Assay methods

Calibration - Measurement function:

Add a number of known standards to the matrix: $X_1, X_2, \ldots, X_n$

Obtain corresponding measurements: $y_1, y_2, \ldots, y_n$

Fit a linear regression (y on X): $y_i = \alpha + \beta \cdot X_i + e_i$

Estimated measurement function: $y = g(X) = \hat{\alpha} + \hat{\beta} \cdot X$

If we observe y_{Obs} then estimate X by: $\hat{X} = g^{-1}(y_{Obs}) = (y_{Obs} - \hat{\alpha}) / \hat{\beta}$

Sometimes called the classical estimator

Assay methods

Calibration - Measurement function:

Why not regress X on y ?: $X_i = \alpha^* + \beta^* \cdot y_i + e_i^*$

where $\alpha^* = \frac{\alpha}{\beta}$ $\beta^* = \frac{1}{\beta}$ $e^* = \frac{1}{\beta} \cdot e$

And then obtain (the prediction): $\hat{X}^* = \hat{\alpha}^* + \hat{\beta}^* \cdot y_{Obs}$

The theory is difficult, but in general: $\hat{X} \neq \hat{X}^*$

The slopes are related: $\hat{\beta} \cdot \hat{\beta}^* = r^2$

If $r^2 = 1$ the observations follows a straight line perfectly, i.e. without any random error. $\hat{\beta}^* = 1/\hat{\beta}$

Thus in general we have: $\hat{\beta}^* < \frac{1}{\hat{\beta}}$

Assay methods

Calibration - Measurement function:

The classical estimator is generally preferred

Confidence interval for $\hat{X}$ can be obtained

Be aware: The notation used in the litterature can differ!

EXAMPLE 6

Calibration - Measurement function:

- insulin bioassay (X) and absorbance readings (y)
- Inverse estimation of X

DATA:

X	y
1	.19
2	.44
3	.63
4	.86
5	1.04

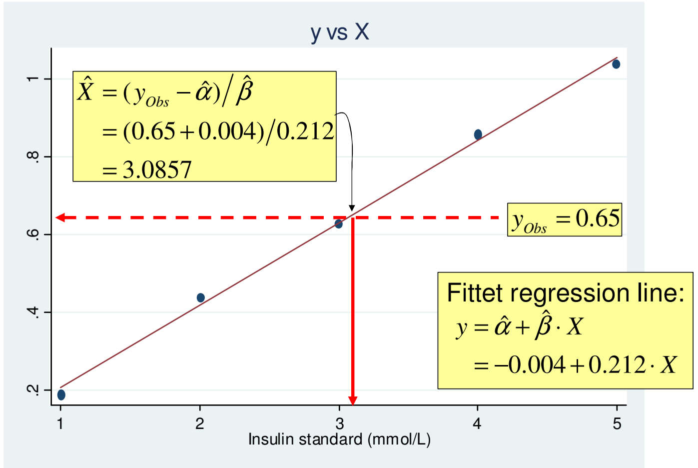
Known standards of insulin (X)

Observed readings of absorbance (y)

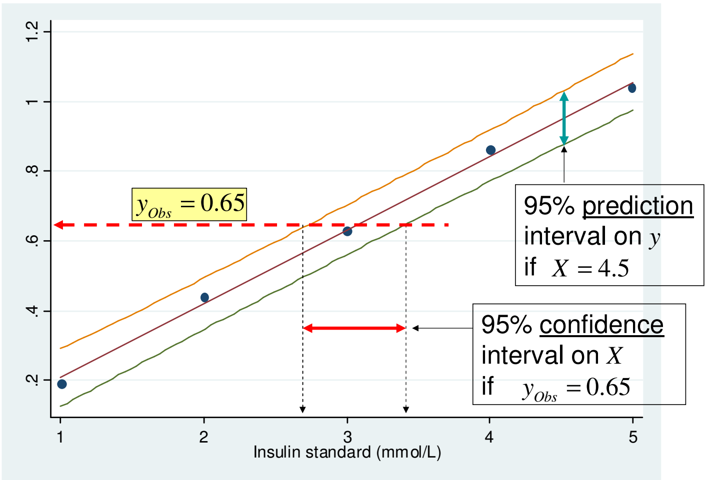
How to estimate insulin concentration $(\hat{X})$ from an observed absorbance reading of $y_{Obs} = 0.65$?

How to obtain a confidence interval for $\hat{X}$?

EXAMPLE 6



EXAMPLE 6



EXAMPLE 6

How to obtain a confidence interval on true value X given an absorbance reading of $y_{Obs} = 0.65$?

Seems simple from figure (slide 12), but we skip the formulaes…

Use ado-file: **regxcii.ado**

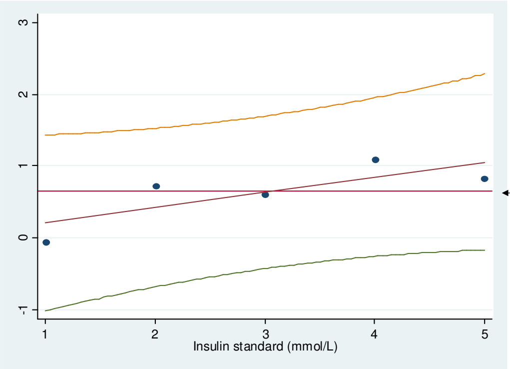
Fit the linear regression of y on X and use **regxcii** as a postestimation command:

```
regress y X       [output skipped]

regxcii 0.65

xhat (CI) = 3.0849057  (2.7489128; 3.4208986)   [Level = 95%]

regxcii 0.70, level(90)

xhat (CI) = 3.3207548  (3.071831; 3.5696785)   [Level = 90%]
```

EXAMPLE 6

What can goes wrong?

If the slope b is badly determined, e.g. not significant!!!



95% CI???

Measurement error models

The next example shows how to compare two assay methods in a design with repeated measurements.

Note: we assume the two methods are "well-calibrated" on known standards.

But how do the methods perform on "real" samples?

Could the methods be influenced by the presence of some "contaminating material" in the sample, e.g. material which could amplify or attenuate the measured "signal"?

In that case: the relationship between the two methods could be Y=a+b*X, where b≠1 and not Y=X !?

EXAMPLE 7

Assay comparison of two methods for measurement of serum gentamicin concentration:

X = Enzyme-mediated immunoassay (EMIT), reference

Y = Fluoro-immunoassay (FIA), new method

56 patients receiving gentamicin, measured twice with each method (separate assays)

"True" or "error free" assay values: $X_i, Y_i, i = 1,2,...,56$

Observed values:

$$x_{i,j} = X_i + u_{i,j},\ u_{i,j} \sim N(0,\sigma_u^2)$$
$$y_{i,j} = Y_i + e_{i,j},\quad e_{i,j} \sim N(0,\sigma_e^2)$$
$$i = 1,2,...,56 \qquad j = 1,2$$

EXAMPLE 7

A model that links Y and X, and the observed values:

Observed values →

$$x_{i,j} = X_i + u_{i,j},\ u_{i,j} \sim N(0,\sigma_u^2)$$
$$y_{i,j} = Y_i + e_{i,j},\quad e_{i,j} \sim N(0,\sigma_e^2)$$
$$Y_i = \alpha + \beta \cdot X_i + q,\ q_i \sim N(0,\sigma_q^2)$$

← Independent errors

$H: \alpha = 0, \beta = 1?$

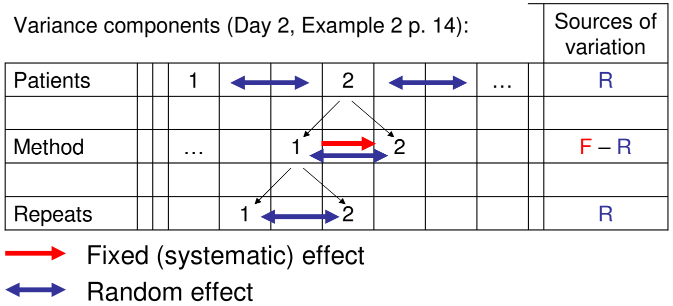
Regression model between "error-free" assay values

This is a structural equation model.
Not fully implemented in Stata. (SAS: PROC TCALIS)

Difficult to estimate α and β.

EXAMPLE 7

Variance components (Day 2, Example 2 p. 14):



We can fit a mixed model (variance component model) equivalent to the model on p. 16

EXAMPLE 7

Use **loneway** to obatin inter- and intraindividual variation:

	sd(inter)	sd(intra)
EMIT	4.47	0.907
FIA	4.19	0.441

Set up **xtmixed** to fit the model on p. 16:

```
xi: xtmixed x i.emit || ///
  ptnr:  emit fia, nocons cov(un) ///
  residuals(independent, by(emit))
```

Note: Exactly the same sd's as above (intra – residual)!

Calculate ICC: $$ICC(FIA) = \frac{4.19^2}{4.19^2 + 0.441^2} = 0.989$$

EXAMPLE 7

Modify **xtmixed** to fit the model on p. 24 with $\beta = 1$:

```
xi: xtmixed x i.emit || ///
     ptnr: emit, ///
     residuals(independent, by(emit))
```

Extract of output:

	sd(inter)	sd(emit)	sd(intra)
EMIT	4.203	0.943	0.907
FIA	4.203	0	0.439

Close to previous values

Note: sd(inter EMIT) = $\sqrt{4.203^2 + 0.943^2} = 4.307$

Close to – but smaller than previous model (4.47)

EXAMPLE 7

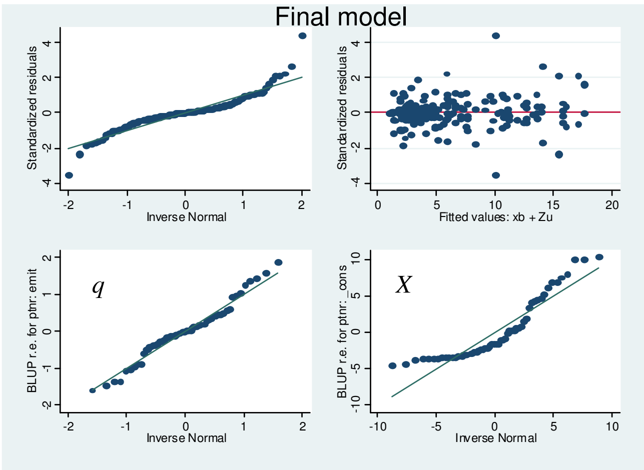
Likelihood ratio test (beta=1): p=0.256. Accept.

What are the assumptions?

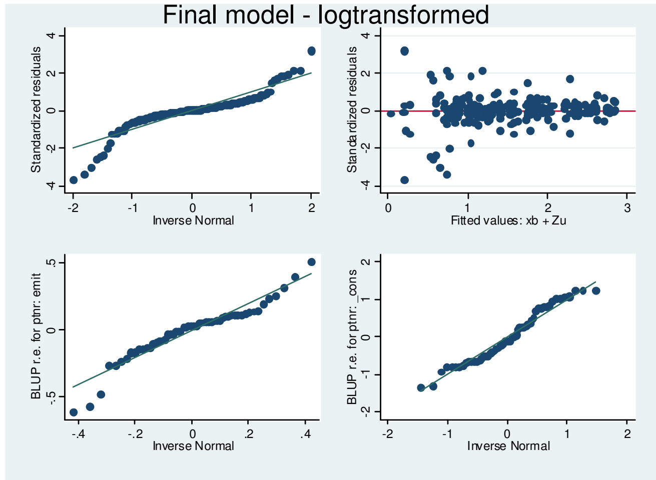
The measurement errors:
- Normal distribution for each method
- Independent (also between methods)

The individual levels (X's):
- Normal distributed
- Regression errors (q's) normal distributed
- Independent (also wrt. measurement errors)

EXAMPLE 7



EXAMPLE 7



EXAMPLE 7

Final model: Construction of limits of agreement (LoA)

One measurement with each method:

Note that:

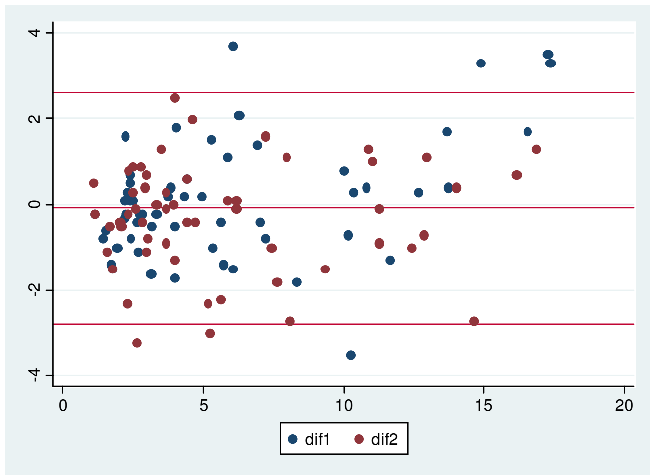
$$
\begin{aligned}
y - x &= (Y + e) - (X + u) \\
&= ([\alpha + \beta \cdot X + q] + e) - (X + u) \\
&= (\alpha + X + q + e) - (X + u) \qquad (\beta = 1) \\
&= \alpha + q + e - u
\end{aligned}
$$

Hence:

$$
\begin{aligned}
sd(\text{dif}) &= sd(y - x) = \sqrt{sd^2(q) + sd^2(e) + sd^2(u)} \\
&= \sqrt{0.943^2 + 0.907^2 + 0.439^2} \\
&= 1.38
\end{aligned}
$$

LoA: $-0.0839 \pm 1.96 \cdot 1.38 = (-2.79; 2.62)$

EXAMPLE 7



Measurement error models

Special designs

The repeated measurements shall represent ”a new measurement”, i.e. using a new assay with re-calibration, and not duplicate measurements in the same assay. Otherwise the estimated measurement error reflects intra-assay random variation.

Be aware of matrix effects in assay methods.

The designs can be complicated and so the analysis.

Consider to ask a statistician in advance!