Compétence et courbe d’apprentissage

Dr David Biau

MCU-PH, Chirurgie orthopédique, Cochin

February 6, 2015
Plan

1. Background
   - Courbe d'apprentissage?
   - Performance?
   - Compétence?

2. LC-CUSUM test
   - Statistical definition
   - Représentation graphique
   - Worked example

3. Applications
   - Competency of radiology residents
   - Competency of gynecology residents
   - Competency of gynecology residents (2)
   - Competency of anesthesiology residents

4. Discussion
Background
T. Wright

“Repetition of the same operation results in less time or effort expended on that operation”

1

courbe d’apprentissage: représentation graphique de la performance au cours de la phase d’apprentissage

1Wright. Factors Affecting the Cost of Airplanes. J. Aero. Sci. 1936
Courbe d’apprentissage d’un endoscopiste pour des cholangiopancreatographies rétrogrades.\textsuperscript{2}

\textsuperscript{2}Schlup et al. ERCP: a review of technical competency and workload in a small unit. \textit{Gastro Endosc} 1997
Performance?

- courbe d’apprentissage ≠ performance
- performance ⇛ % procédures échouées
- apprentissage ⇛ amélioration de la performance
Performance de l’endoscopiste

Graphique de la somme cumulative des échecs

Cumulative sum of failures

Observations

0 50 100 150 200 250 300 350 400 450 500 550
0
10
20
30
40
50
60

4%

18%
Compétence

Être compétent c’est avoir atteint un niveau de performance prédéfini

Compétence déclarée par

- observation d’un tuteur
- réalisation d’un certain nombre de procédure
- fin d’un cursus (internat)

---


Quand est-ce que la compétence est atteinte?
LC-CUSUM test$^5$

Learning-Curve CUmulative SUMmation test.

---

$^5$Biau, Porcher. Monitoring a process from an out of control to an in control state. Application to the learning curve. *Statistics in Medicine*, in press.
Hypotheses

- H0 “performance est inadéquate”
- H1 “performance est adéquate”

⇒ LC-CUSUM teste *séquentiellement* H0 contre H1
Définition

**LC-CUSUM score**

\[ S_t = \max(0, S_{t-1} + W_t), \quad S_0 = 0 \]

- **\( S_t \):** LC-CUSUM score au temps/obs \( t \)
- **\( W_t \):** sample weight
- **\( S_t \geq h \iff \) performance adéquate démontrée**
Sample weight

\[ W_t = \log \left[ \frac{p_0^{X_t}(1 - p_0)^{1-X_t}}{(p_0 + \delta)^{X_t}(1 - p_0 - \delta)^{1-X_t}} \right] \]

- \( X_t \): 0 pour un succès et 1 pour un échec
- \( p_0 \): niveau de performance adéquate
- \( \delta \): zone d’équivalence
Exemple de l’ERCP

- $p_0 = 10\%$ (performance adéquate)
- $\delta = 7.5\%$ (équivalence)
- Sample weight:
  - $W_t$ pour un succès = $+0.09$
  - $W_t$ pour un échec = $-0.56$
Outcome: 5S, 1F

adequate perf. (H1)

continuation region

0 = absorbing barrier

h = limit
Outcome: 5S, 2F, 5S, 8F

- adequates perf. (H1)
- h = limit
- continuation region
- 0 = absorbing barrier

LC−CUSUM test

Applications

Discussion
Représentation graphique

Outcome: 5S, 2F, 5S, 8F, 9S, 1F, 10S

- LC-CUSUM score
- Time/Observations
- h = limit
- 0 = absorbing barrier
- adequate perf. (H1)
- continuation region

0 = absorbing barrier
Performance du test

- Erreurs de type I et de type II\(^6\)
  - erreur type I: 100%
  - erreur type II: 0%; puissance 100%

- ADR (Alarm Discovery Rate) pour un nombre d’observation, sous H0 et sous H1
  - ADR sous H1: prob. de déclencher une alarme lorsque l’étudiant est compétent
  - ADR sous H0: prob. de déclencher une alarme lorsque l’étudiant n’est pas compétent

\(^6\)Neyman J., Pearson E. On the problem of the most efficient tests of statistical hypotheses. *Philos. Trans. R. Soc. Lond. A.*, 1933
Competency of the endoscopist

- Process: ERCP
- Outcome: successful cannulation of the bile duct
- Adequate performance: 10% failure; equivalence: 7.5% 
- \( h = 4.47 \)
  - for 500 procedures
  - ADR1 = 99%
  - ADR0 = 10%
LC-CUSUM test for the endoscopist

Applications
Transvaginal sonography

- Process: diagnosis of endometriomas by transvaginal sonography
- Outcome: identical findings as senior sonographer
- Adequate performance: 15% failure ($\delta = 7.5\%$)
- $h = 1 \iff$
  - for 75 procedures
  - ADR1 = 89%
  - ADR0 = 10%
LC-CUSUM test

Applications

Discussion

Observations

LC−CUSUM score

0 5 10 15 20 25 30 35 40 45

0 0.2 0.4 0.6 0.8 1

Trainee 1

Trainee 2

Trainee 3

Trainee 4

8

Bazot et al. Learning curve of transvaginal ultrasound for the diagnosis of endometriomas assessed by the LC-CUSUM. Fertil Steril. 2011
Embryo transfer

- Process: embryo transfer
- Outcome: positive hCG test
- Adequate performance: 60% failure ($\delta = 10\%$)
- $h = 1.86 \Rightarrow$
  - for 100 procedures
  - ADR1 = 90%
  - ADR0 = 1.7%
Observations
LC−CUSUM score
0 20 40 60 80 100
0.0 0.5 1.0 1.5 2.0
Trainee 1
Trainee 2
Trainee 3
Trainee 4
Trainee 5

Dessolle L et al. How soon can I be proficient in embryo transfer. Human Reprod, 2010
Fetoscopic laser photocoagulation

- process: selective laser photocoagulation in Twin-Twin Transfusion Syndrome
- outcome: at least one survivor
- adequate performance: 18% failure
- $h = 0.95$
  - for 75 procedures
  - ADR1 = 85%
  - ADR0 = 20%
Process: TEA
Outcome:
Adequate performance: 60% failure ($\delta = 10\%$)
$h = 1.86$
\begin{itemize}
  \item for 100 procedures
  \item ADR1 = 90%
  \item ADR0 = 1.7%
\end{itemize}
Weil G et al. Learning of specific procedures and residents technical skills: assessment using a statistical process control method. In progress
Discussion
Mise en place du test LC-CUSUM

- Processus: clairement défini ("Repetition of the same operation")
- Outcome: mesure effectivement le processus (infection et PTH)
- Niveau de performance
- n(obs) autorisé, limite h
- ADR sous H1 et H0
Choisir $h$ et $n(\text{obs})$

définis par l’utilisateur
pas de solution analytique
simulation ($\approx 10,000$ samples)
ADR1 et ADR0
LC−CUSUM test

Applications

Discussion

Time/Observations

LC−CUSUM score

0 10 20 30 40 50 60 70 80

0 1 2 3

\( h = 0.75 \)

power = 99%

\( P(\text{false/\text{al.}}) = 46\% \)

10% failure rate

30% failure rate

10% failure rate

30% failure rate

10% failure rate

30% failure rate

10% failure rate

30% failure rate

10% failure rate

30% failure rate

10% failure rate

30% failure rate

10% failure rate

30% failure rate

10% failure rate

30% failure rate

10% failure rate

30% failure rate

10% failure rate

30% failure rate

10% failure rate

30% failure rate

10% failure rate

30% failure rate

10% failure rate

30% failure rate

10% failure rate

30% failure rate

10% failure rate

30% failure rate

10% failure rate

30% failure rate

10% failure rate

30% failure rate

10% failure rate

30% failure rate

10% failure rate

30% failure rate

10% failure rate

30% failure rate

10% failure rate

30% failure rate

10% failure rate

30% failure rate

10% failure rate

30% failure rate

10% failure rate

30% failure rate

10% failure rate

30% failure rate

10% failure rate

30% failure rate

10% failure rate

30% failure rate

10% failure rate

30% failure rate

10% failure rate

30% failure rate

10% failure rate

30% failure rate

10% failure rate

30% failure rate

10% failure rate

30% failure rate

10% failure rate

30% failure rate

10% failure rate

30% failure rate

10% failure rate

30% failure rate

10% failure rate

30% failure rate

10% failure rate

30% failure rate

10% failure rate

30% failure rate

10% failure rate

30% failure rate

10% failure rate

30% failure rate

10% failure rate

30% failure rate

10% failure rate

30% failure rate

10% failure rate

30% failure rate

10% failure rate

30% failure rate

10% failure rate

30% failure rate

10% failure rate

30% failure rate

10% failure rate

30% failure rate

10% failure rate

30% failure rate

10% failure rate

30% failure rate

10% failure rate

30% failure rate

10% failure rate

30% failure rate

10% failure rate

30% failure rate

10% failure rate

30% failure rate

10% failure rate

30% failure rate

10% failure rate

30% failure rate

10% failure rate

30% failure rate

10% failure rate

30% failure rate

10% failure rate

30% failure rate

10% failure rate

30% failure rate

10% failure rate

30% failure rate

10% failure rate

30% failure rate

10% failure rate

30% failure rate

10% failure rate

30% failure rate

10% failure rate

30% failure rate

10% failure rate

30% failure rate

10% failure rate

30% failure rate

10% failure rate

30% failure rate

10% failure rate

30% failure rate

10% failure rate

30% failure rate

10% failure rate

30% failure rate

10% failure rate

30% failure rate

10% failure rate

30% failure rate

10% failure rate

30% failure rate

10% failure rate

30% failure rate

10% failure rate

30% failure rate

10% failure rate

30% failure rate

10% failure rate

30% failure rate

10% failure rate

30% failure rate

10% failure rate

30% failure rate

10% failure rate

30% failure rate

10% failure rate

30% failure rate

10% failure rate

30% failure rate

10% failure rate

30% failure rate

10% failure rate

30% failure rate

10% failure rate

30% failure rate

10% failure rate

30% failure rate

10% failure rate

30% failure rate

10% failure rate

30% failure rate

10% failure rate

30% failure rate

10% failure rate

30% failure rate

10% failure rate

30% failure rate

10% failure rate

30% failure rate

10% failure rate

30% failure rate

10% failure rate

30% failure rate

10% failure rate

30% failure rate

10% failure rate

30% failure rate

10% failure rate

30% failure rate

10% failure rate

30% failure rate

10% failure rate

30% failure rate

10% failure rate

30% failure rate

10% failure rate

30% failure rate

10% failure rate

30% failure rate

10% failure rate

30% failure rate

10% failure rate

30% failure rate

10% failure rate

30% failure rate

10% failure rate

30% failure rate

10% failure rate

30% failure rate

10% failure rate

30% failure rate

10% failure rate

30% failure rate

10% failure rate

30% failure rate

10% failure rate

30% failure rate

10% failure rate

30% failure rate

10% failure rate

30% failure rate

10% failure rate

30% failure rate

10% failure rate

30% failure rate

10% failure rate

30% failure rate

10% failure rate

30% failure rate

10% failure rate

30% failure rate

10% failure rate

30% failure rate

10% failure rate

30% failure rate

10% failure rate

30% failure rate

10% failure rate

30% failure rate

10% failure rate

30% failure rate

10% failure rate

30% failure rate

10% failure rate

30% failure rate

10% failure rate

30% failure rate

10% failure rate

30% failure rate

10% failure rate

30% failure rate

10% failure rate

30% failure rate

10% failure rate

30% failure rate

10% failure rate

30% failure rate

10% failure rate

30% failure rate

10% failure rate

30% failure rate

10% failure rate

30% failure rate

10% failure rate

30% failure rate

10% failure rate

30% failure rate

10% failure rate

30% failure rate

10% failure rate

30% failure rate

10% failure rate

30% failure rate

10% failure rate

30% failure rate

10% failure rate

30% failure rate

10% failure rate

30% failure rate

10% failure rate

30% failure rate

10% failure rate

30% failure rate

10% failure rate

30% failure rate

10% failure rate

30% failure rate

10% failure rate

30% failure rate
LC-CUSUM test

Applications

Discussion

Time/Observations

LC−CUSUM score

0 10 20 30 40 50 60 70 80

0 1 2 3

h = 2

10% failure rate

30% failure rate

h

power= 52%

P(false/al.)= 3%
LC−CUSUM score

Time/Observations

h = 2.5

power = 76%

P(false/αl.) = 6%

10% failure rate

30% failure rate
Fin de la surveillance?

1. Compétence est démontrée
   ➞ arrêter la surveillance?
   ➞ utiliser un test pour s’assurer que le processus reste contrôler

2. Compétence non démontrée
   ➞ on déclare l’étudiant non compétent?
   ➞ on recommence le test? (mais on ne contrôle plus les risques d’erreurs)
   ➞ envisager cette possibilité au début
test LC-CUSUM
Détermine quand un étudiant est compétent

- ARRÊTER de former ce qui n’en ont plus besoin
- CONTINUER à former ce qui en ont encore besoin
- utiliser les ressources de manière rationnelle
Compétence et courbe d’apprentissage

Dr David Biau

MCU-PH, Chirurgie orthopédique, Cochin

February 6, 2015
## Example de simulation pour h

<table>
<thead>
<tr>
<th>h</th>
<th>18%</th>
<th>23%</th>
<th>27%</th>
<th>30%</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.5</td>
<td>99.4</td>
<td>96.0</td>
<td>88.6</td>
<td>80.6</td>
</tr>
<tr>
<td>0.6</td>
<td>96.6</td>
<td>86.0</td>
<td>70.2</td>
<td>56.3</td>
</tr>
<tr>
<td>0.7</td>
<td>93.2</td>
<td>74.5</td>
<td>53.1</td>
<td>38.8</td>
</tr>
<tr>
<td>0.75</td>
<td>91.7</td>
<td>71.7</td>
<td>50.3</td>
<td>35.9</td>
</tr>
<tr>
<td>0.85</td>
<td>84.6</td>
<td>58.2</td>
<td>34.7</td>
<td>22.1</td>
</tr>
<tr>
<td>0.9</td>
<td>81.3</td>
<td>52.7</td>
<td>30.1</td>
<td>17.7</td>
</tr>
<tr>
<td>1</td>
<td>75.2</td>
<td>44.3</td>
<td>23.1</td>
<td>12.7</td>
</tr>
<tr>
<td>1.1</td>
<td>67.7</td>
<td>33.6</td>
<td>16.1</td>
<td>7.8</td>
</tr>
<tr>
<td>1.2</td>
<td>60.3</td>
<td>28.5</td>
<td>11.0</td>
<td>5.6</td>
</tr>
<tr>
<td>1.3</td>
<td>54.3</td>
<td>21.1</td>
<td>7.9</td>
<td>3.0</td>
</tr>
<tr>
<td>1.4</td>
<td>46.5</td>
<td>16.5</td>
<td>5.6</td>
<td>1.9</td>
</tr>
<tr>
<td>1.5</td>
<td>41.3</td>
<td>13.0</td>
<td>4.2</td>
<td>1.5</td>
</tr>
</tbody>
</table>