

Pourquoi l'intervalle de confiance standard de Wald pour une proportion est-il toujours utilisé?

Jean-François Petiot*

Jean-Christophe Turlot**

* Laboratoire de Mathématiques de Bretagne Atlantique (LMBA) – UMR 6205 - IUT de Vannes - Université de Bretagne Sud – Université Européenne de Bretagne

** Laboratoire de Mathématiques et de leurs Applications – UMR CNRS 5142- IUT de Pau – Université de Pau et des Pays de l'Adour

Méthode de Wald

$$Z = \frac{\sqrt{n}(\hat{p} - p)}{\sqrt{p(1-p)}} \rightarrow N(0,1) \quad (TCL)$$

$$IC_W(p, 1-\alpha) = \hat{p} \pm u_{1-\alpha/2} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

Conditions de validité ?

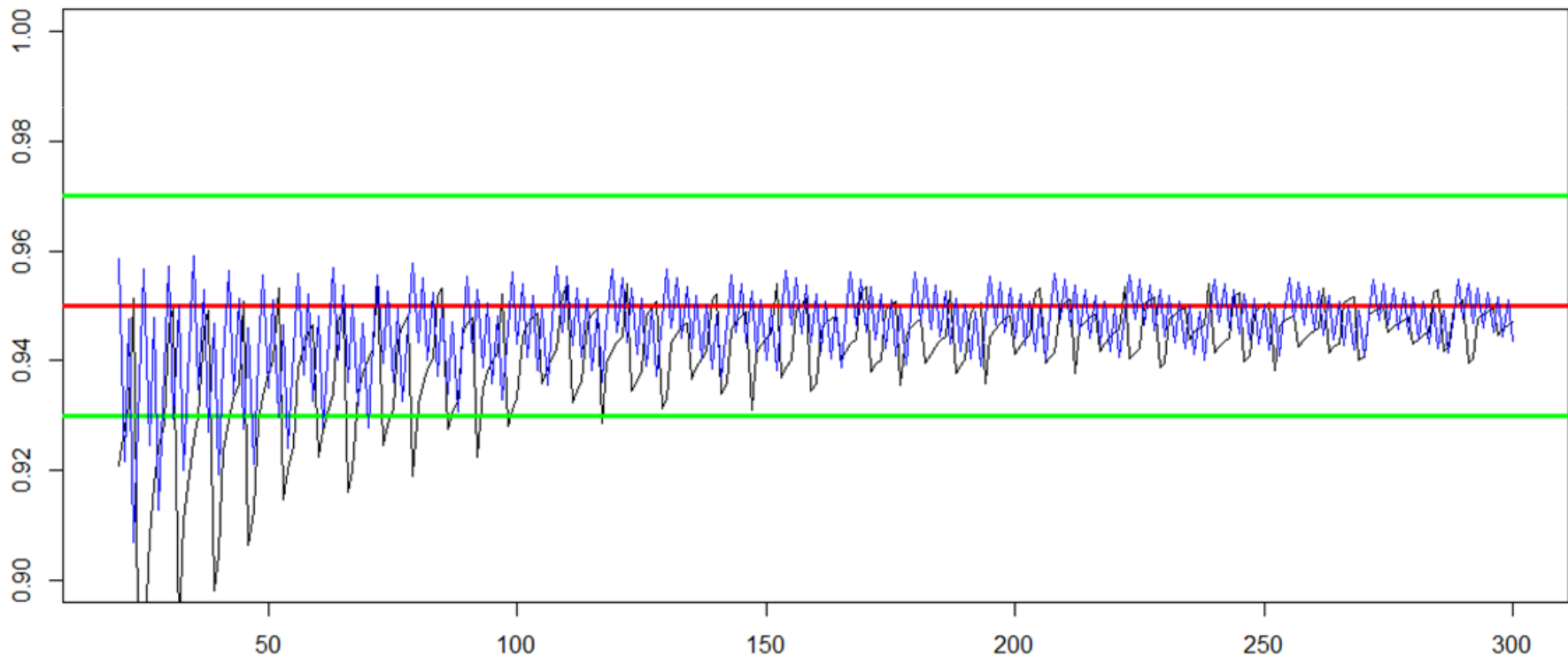
$n\hat{p} \geq 5$ ou 10 et $n(1-\hat{p}) \geq 5$ ou 10 ?

$n \geq 50$ sauf si p est "très petit" ?

etc...

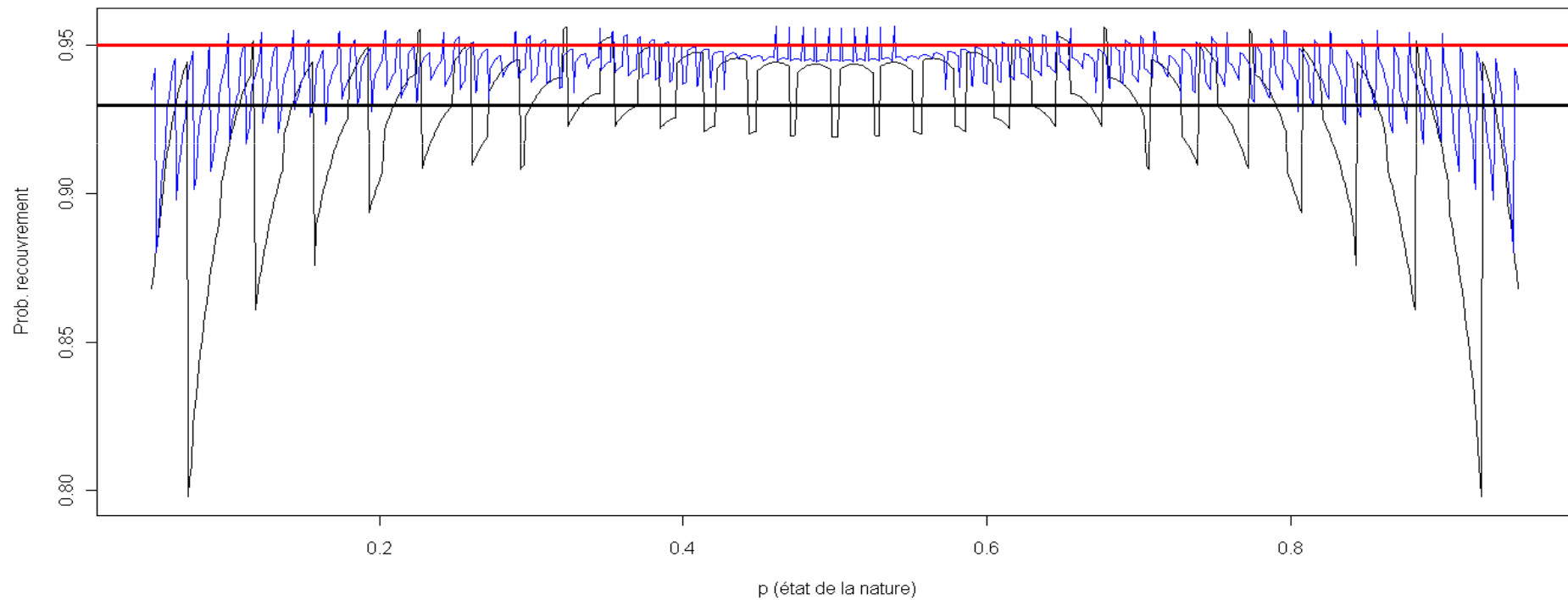
Wald : probabilités de recouvrement de $p=0,2$ (noir) et $p=0,5$ (bleu) n de 20 à 300

Wald standard: Probabilités de recouvrement pour $p=0,2$ (noir) et $p=0,5$ (bleu)



Probabilités de recouvrement de la vraie valeur de p avec $n=40$ et 120

IC WALD standard: Probabilités de recouvrement pour $n=40$ et $n=120$ (en bleu) en fonction de p

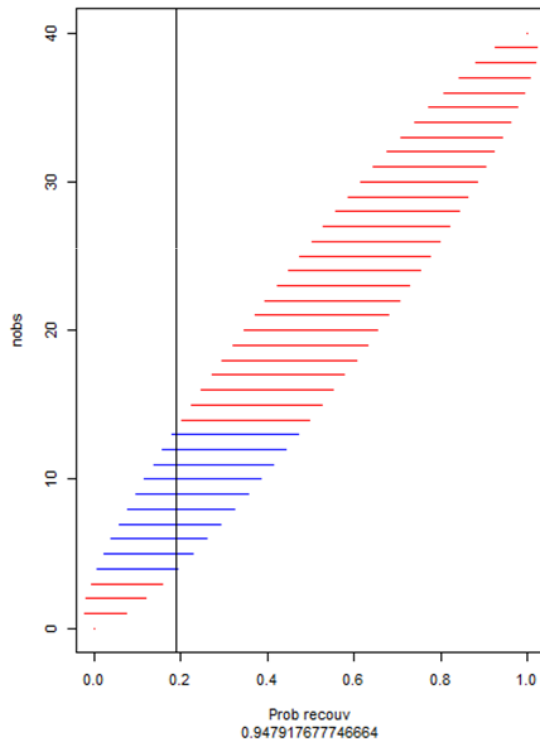


Des probabilités de recouvrement de la vraie valeur de p (ici $p=0,5$) erratiques

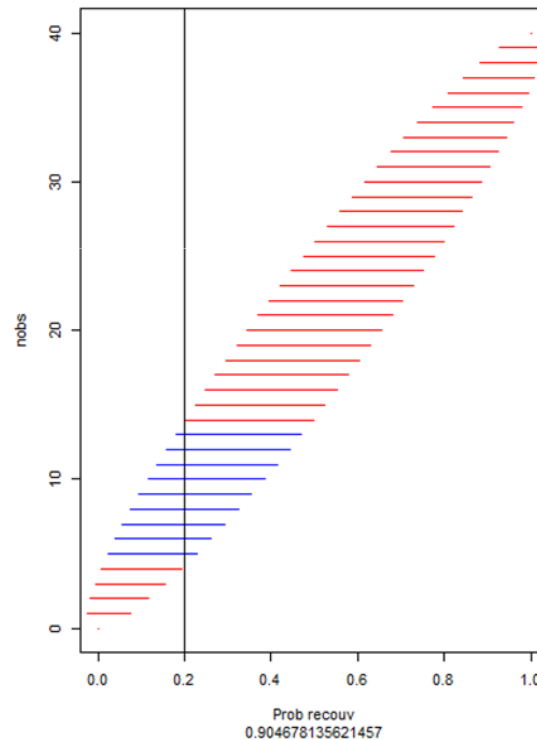
(extrait de Brown, Cai et Dasgupta, Statistical Science, 2001)

| n chanceux | 17 | 20 | 25 | 30 | 35 | 37 | 42 | 44 | 49 |
|-------------------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| $P(0,5 \in IC_W)$ | 0,951 | 0,959 | 0,957 | 0,957 | 0,959 | 0,953 | 0,956 | 0,951 | 0,956 |
| n non chanceux | 10 | 12 | 13 | 15 | 18 | 23 | 28 | 33 | 40 |
| $P(0,5 \in IC_W)$ | 0,891 | 0,854 | 0,908 | 0,882 | 0,904 | 0,907 | 0,913 | 0,920 | 0,919 |

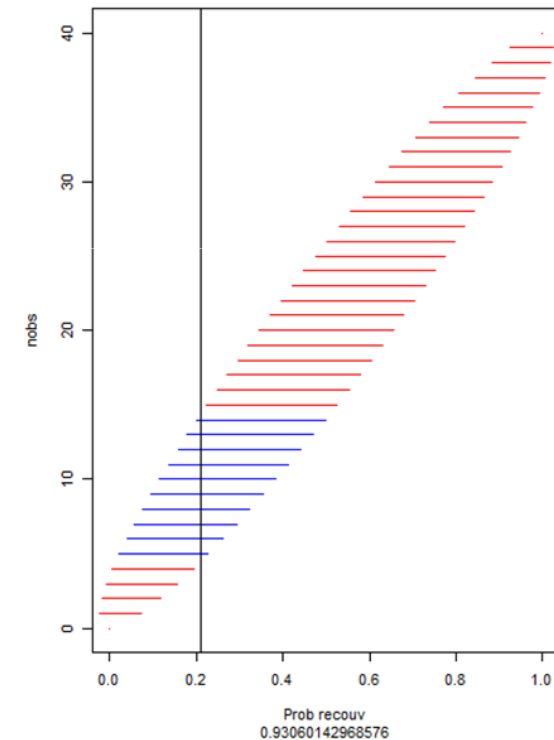
« Zoom » sur les fluctuations de probabilité de recouvrement (méthode : Wald) pour $p=0,19$, $0,20$ et $0,21$ avec $n=40$



0,948



0,905



0,931

Des méthodes alternatives

- Méthodes dites « exactes »
 - Clopper – Pearson (CP)
 - Mid – P
 - Approche bayésienne (loi a priori de Jeffreys)
 - X+U (Stevens)
- Méthodes par approximation normale
 - Wilson (score)
 - Agresti – Coull (AC)
 - Rapport de vraisemblance (RV)
 - Delta-logit
 -

Notations

$(X_i)_{i=1,\dots,n}$ *Bernoulli indépendantes*

$$T = \sum_{i=1}^n X_i \qquad \hat{p} = \frac{T}{n}$$

- Probabilité nominale de recouvrement : $1-\alpha$
- Probabilité réelle de recouvrement : $P(p \in IC_*)$

Méthode (dite exacte) de Clopper – Pearson

$$p \in IC_{CP} \Leftrightarrow P_p(T \geq t) = \alpha/2 \text{ et } P_p(T \leq t) = \alpha/2$$

- De la relation entre les lois binomiale et beta :

$$P(B(n, p) \leq c) = P(\beta(c + 1, n - c) \geq p)$$

- on déduit les bornes de :

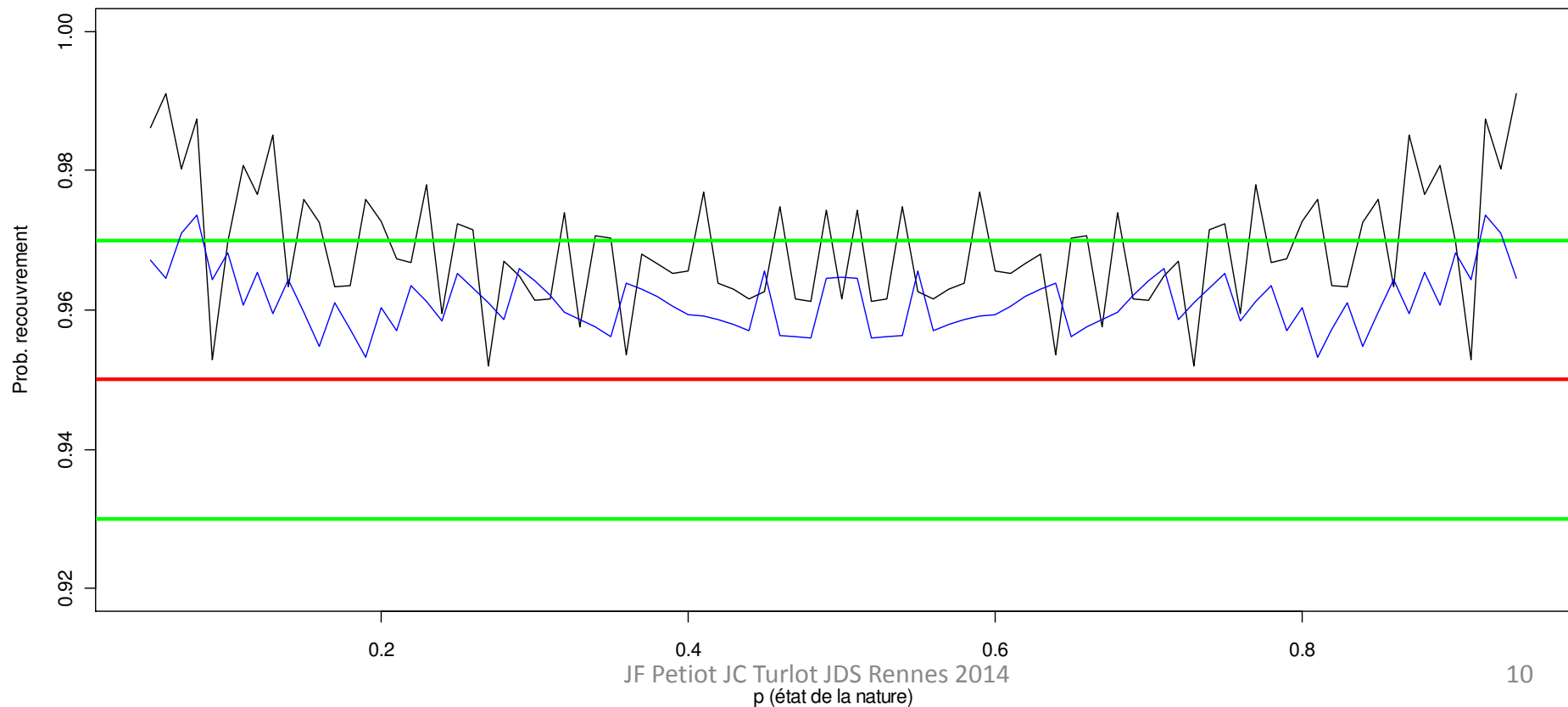
$$l_{CP} = Q\beta(t, n - t + 1, \alpha/2) \quad u_{CP} = Q\beta(t + 1, n - t, 1 - \alpha/2)$$

avec $Q\beta$ quantiles de la loi $\beta(a, b)$

Probabilités de recouvrement de la vraie valeur de p avec n=40 et 120

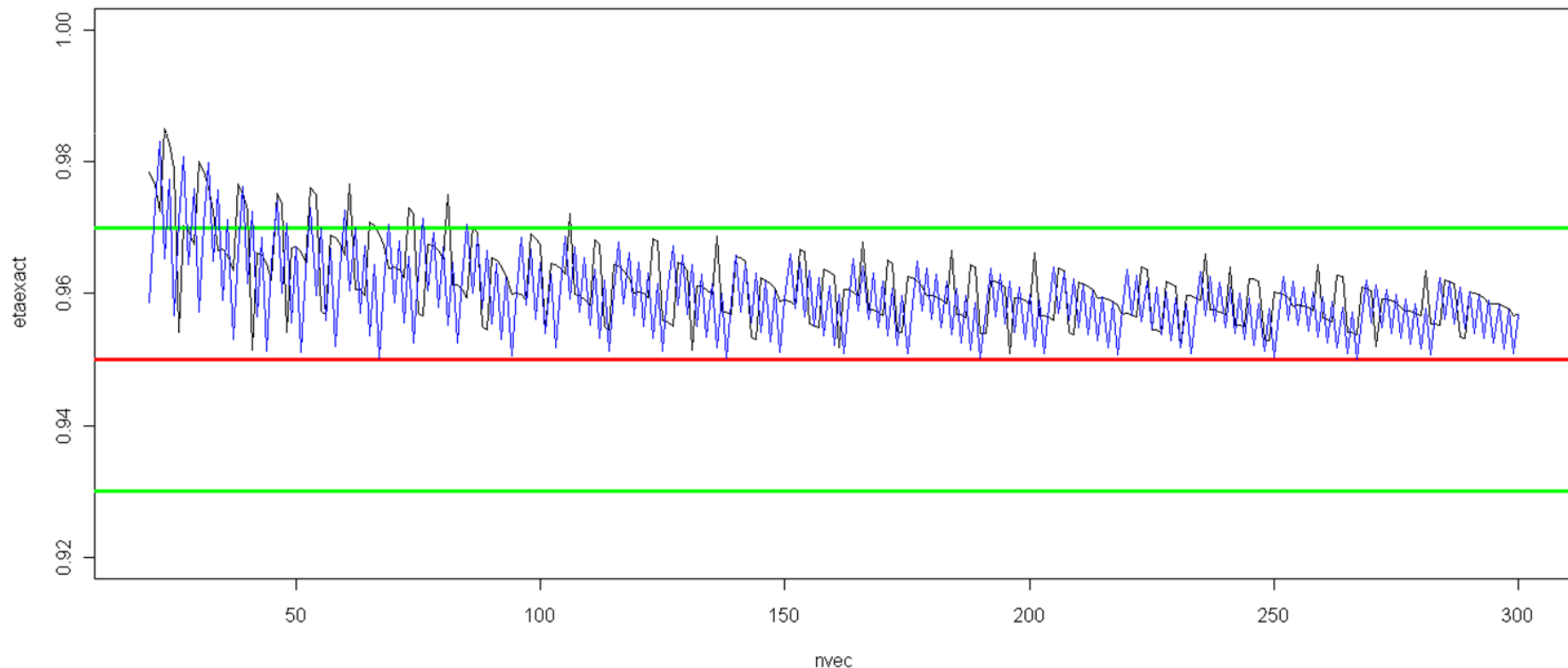
Méthode de Clopper - Pearson

IC Clopper - Pearson : Probabilités de recouvrement en fonction de p avec n=40 (noir) et n=120 (bleu)



Probabilités de recouvrement de la valeur de p : 0,2 (noir) et 0,5 (en bleu) n : de 20 à 300 Méthode de Clopper - Pearson

IC Clopper - Pearson : probabilités de recouvrement de $p=0,2$ (noir) et $p=0,5$ (bleu)



Méthode MID-P

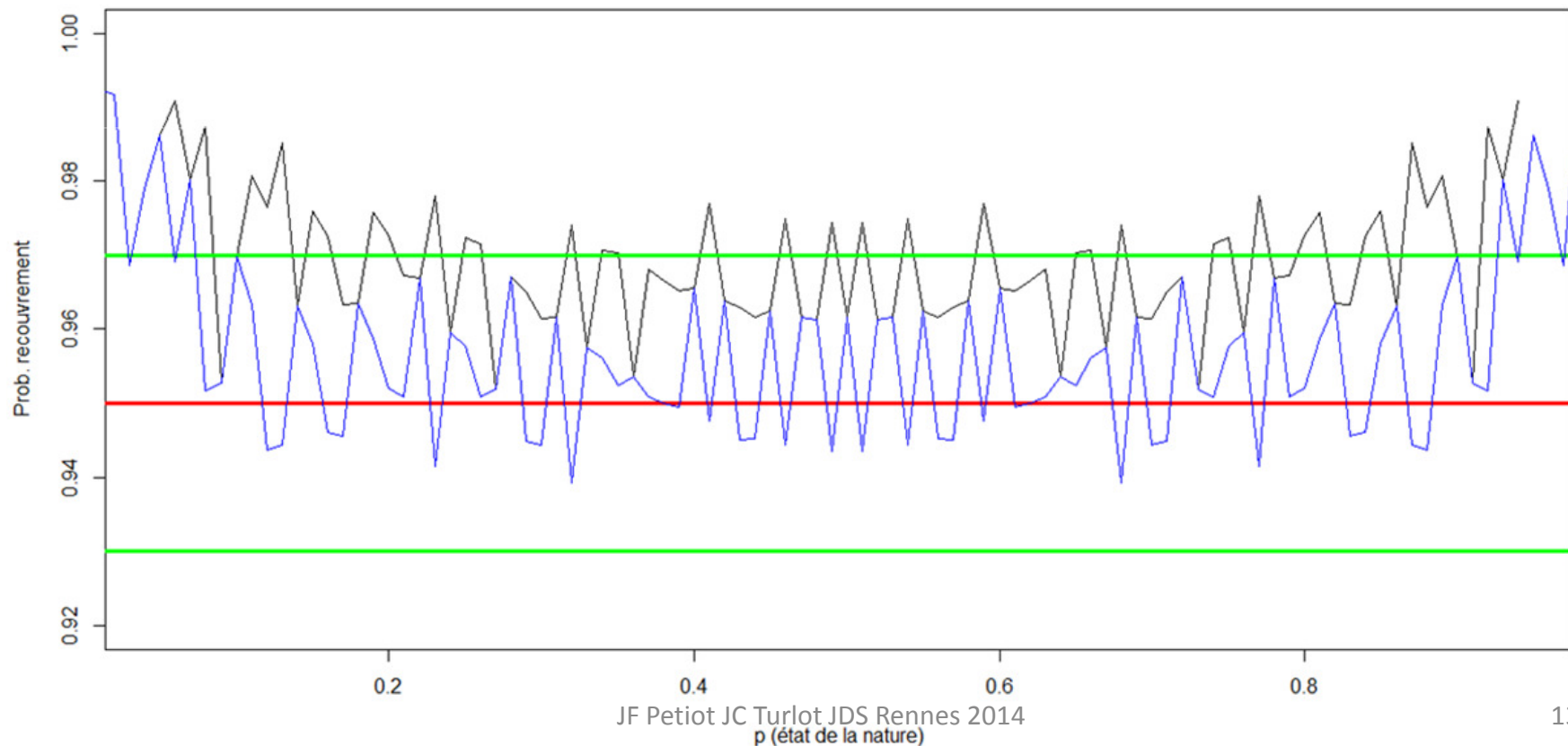
$$p \in IC_{CP} \Leftrightarrow P_p(T > t) + 0,5P(T = t) = \alpha / 2$$

$$\text{et } P_p(T < t) + 0,5P(T = t) = \alpha / 2$$

Les bornes l_{MID-P} et u_{MID-P} n'ont pas d'expression explicite

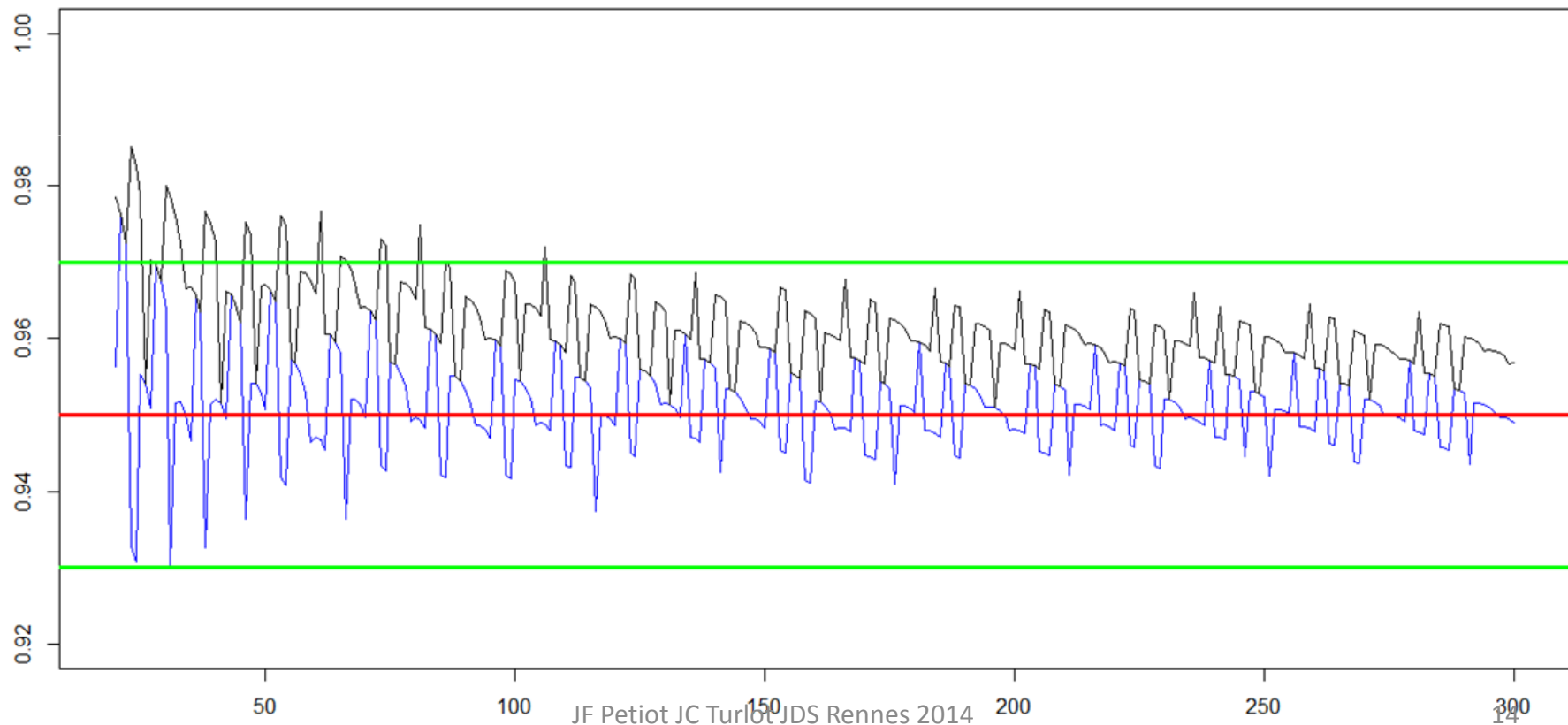
Probabilités de recouvrement de p avec $n=40$: MID-P (en bleu) moins conservative que Clopper – Pearson (noir)

IC Clopper – Pearson et MID-P : Probabilités de recouvrement en fonction de p avec $n=40$ pour CP (noir) et MID-P (bleu)



Probabilités de recouvrement de $p=0,20$ par MID-P (bleu) et Clopper-Pearson (noir) avec n de 20 à 300

CP (noir) et MID-P (bleu): Probabilités de recouvrement de $p=0,20$ en fonction de n



Méthode bayésienne (Jeffreys)

$$T \sim B(n, p)$$

$$\text{Loi a priori: } p \rightarrow \beta(b_1, b_2)$$

$$\text{Loi a posteriori: } p \rightarrow \beta(t + b_1, n - t + b_2)$$

$$l_J(t) = Q\beta(t + b_1, n - t + b_2, \alpha/2)$$

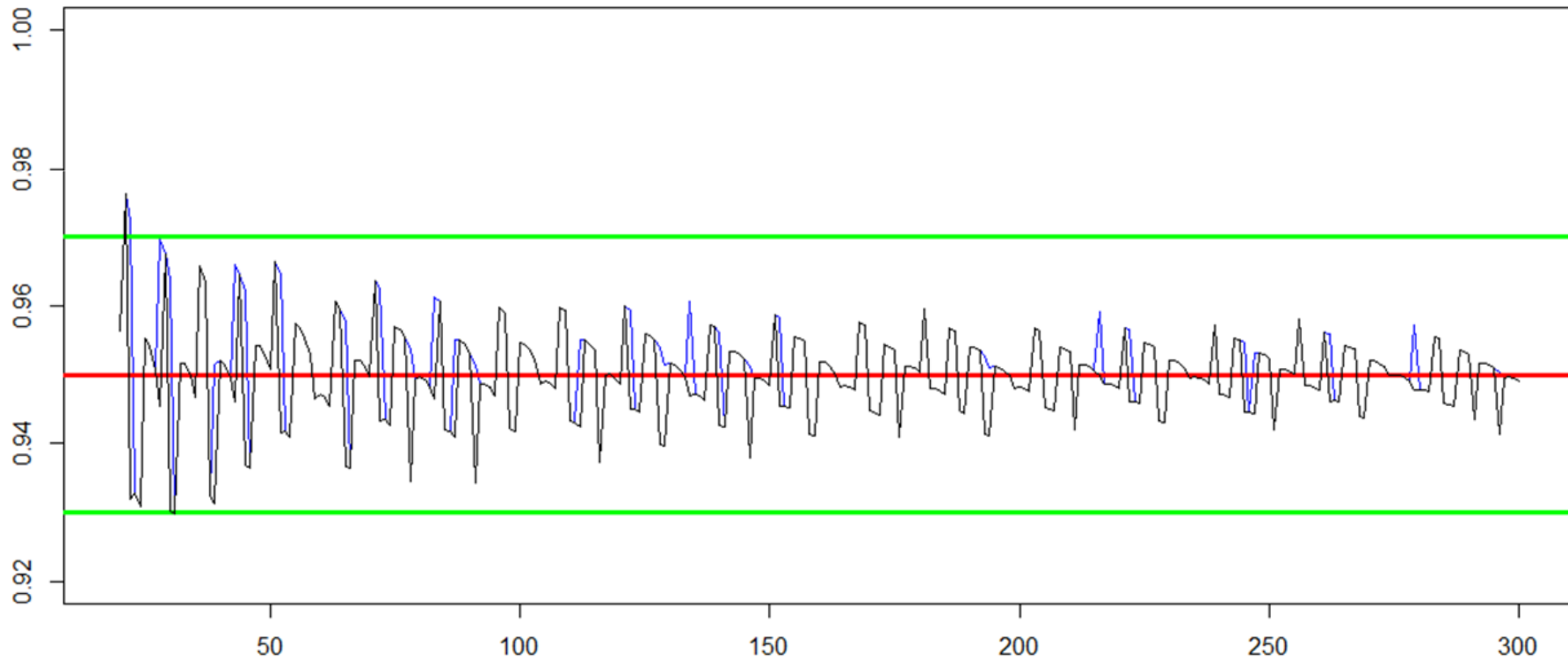
$$u_J(t) = Q\beta(t + b_1, n - t + b_2, 1 - \alpha/2)$$

$$\text{Ici : } (b_1, b_2) = (1/2, 1/2)$$

Loi a priori non informative de Jeffreys

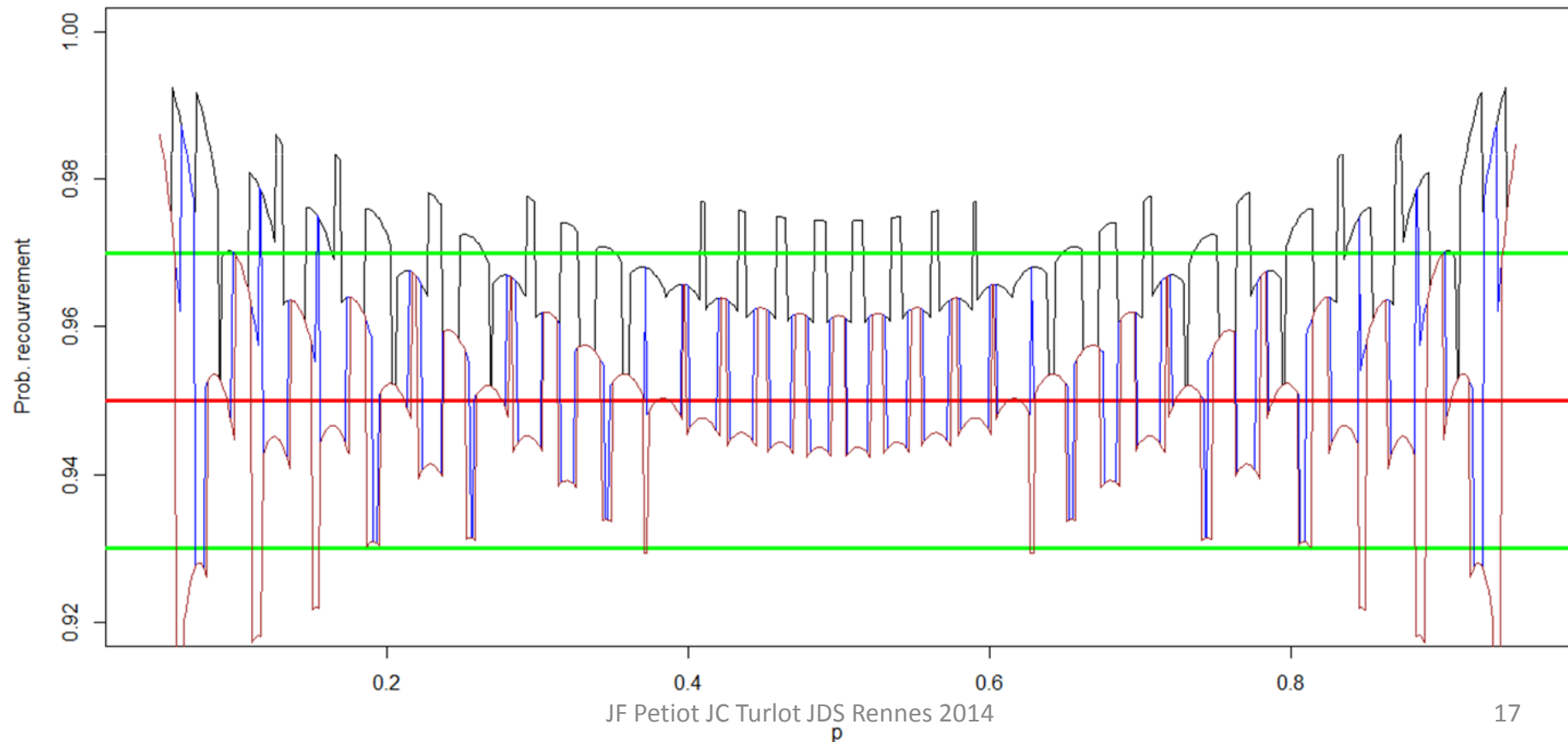
Probabilités de recouvrement de $p=0,20$ avec Jeffreys (noir) et MID-P (bleu) avec n de 20 à 300

MID-P (bleu) et Jeffreys (brun): Probabilités de recouvrement de $p=0,20$ en fonction de n



Probabilités de recouvrement de p par Jeffreys (brun), MID-P (bleu) et Clopper – Pearson (noir) $n=40$

IC CP(noir) MID-P (bleu) et Jeffreys (brun) : Probabilités de recouvrement en fonction de p avec $n=40$



Méthode de Wilson

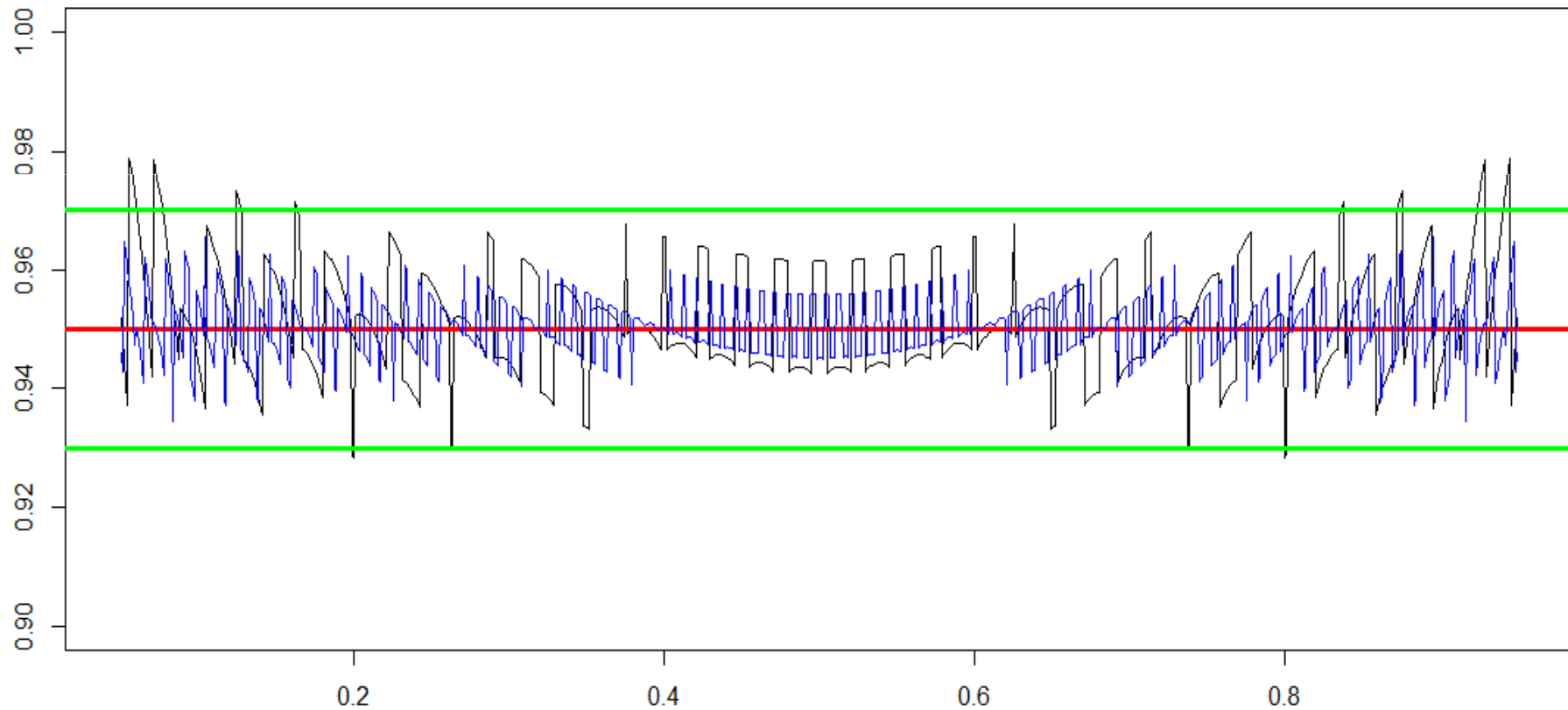
$$Z = \frac{\sqrt{n}(\hat{p} - p)}{\sqrt{p(1-p)}} \rightarrow N(0,1) \quad (TCL)$$

$$\tilde{T} = T + \frac{1}{2}u_{1-\alpha/2}^2 \quad \tilde{n} = n + u_{1-\alpha/2}^2 \quad \tilde{p} = \frac{\tilde{T}}{\tilde{n}}$$

$$IC_{WIL}(p, 1-\alpha) = \tilde{p} \pm u_{1-\alpha/2} \frac{n^{1/2}}{n + u_{1-\alpha/2}^2} \left(\hat{p}(1-\hat{p}) + \frac{u_{1-\alpha/2}^2}{4n} \right)^{1/2}$$

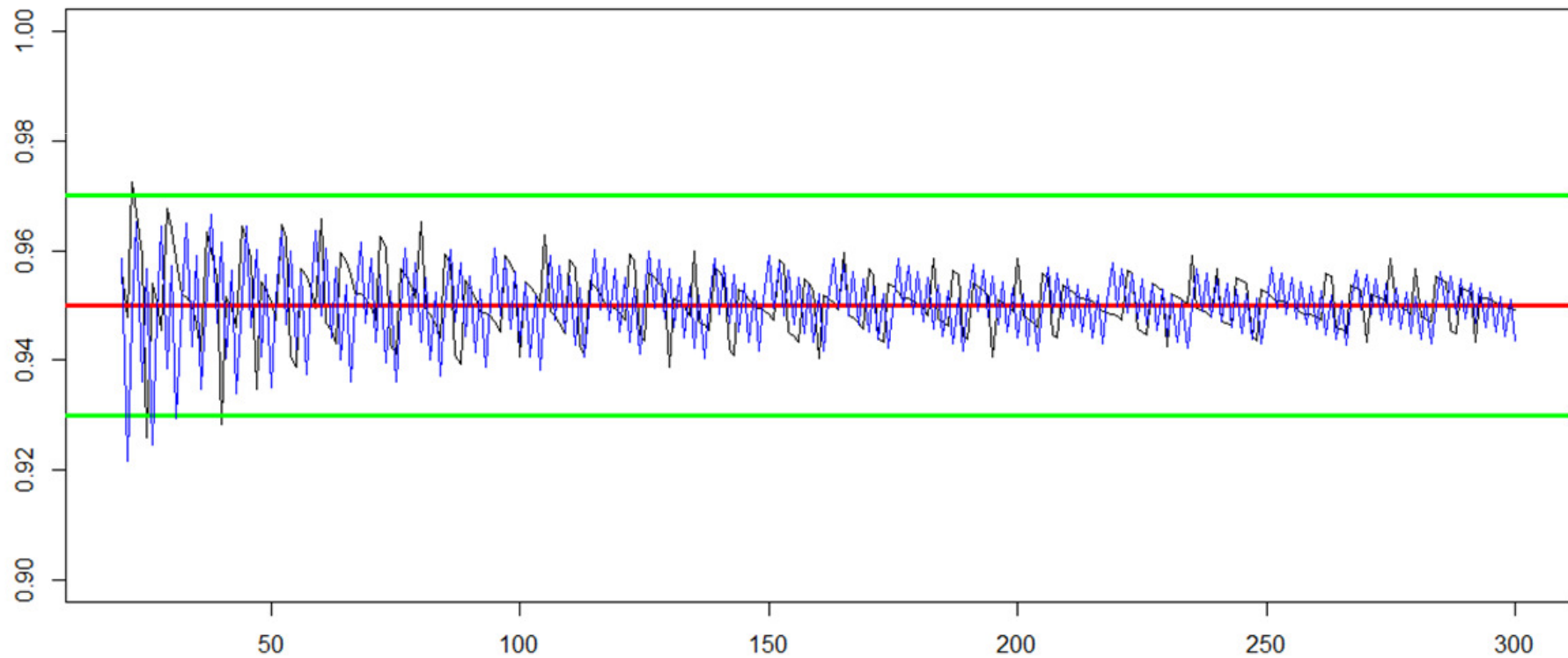
Wilson : probabilités de recouvrement de p avec $n=40$ (noir) et $n=120$ (bleu)

Wilson: Probabilités de recouvrement pour $n=40$ (noir) et $n=120$ (bleu)



Wilson: probabilités de recouvrement de $p=0,2$ (noir) et $p=0,5$ (bleu) n de 20 à 300

Wilson: Probabilités de recouvrement pour $p=0,2$ (noir) et $p=0,5$ (bleu)



Méthode d'Agresti - Coull

$$\tilde{n} = n + u_{1-\alpha/2}^2 \quad \tilde{T} = T + \frac{1}{2}u_{1-\alpha/2}^2 \quad \tilde{p} = \frac{\tilde{T}}{\tilde{n}}$$

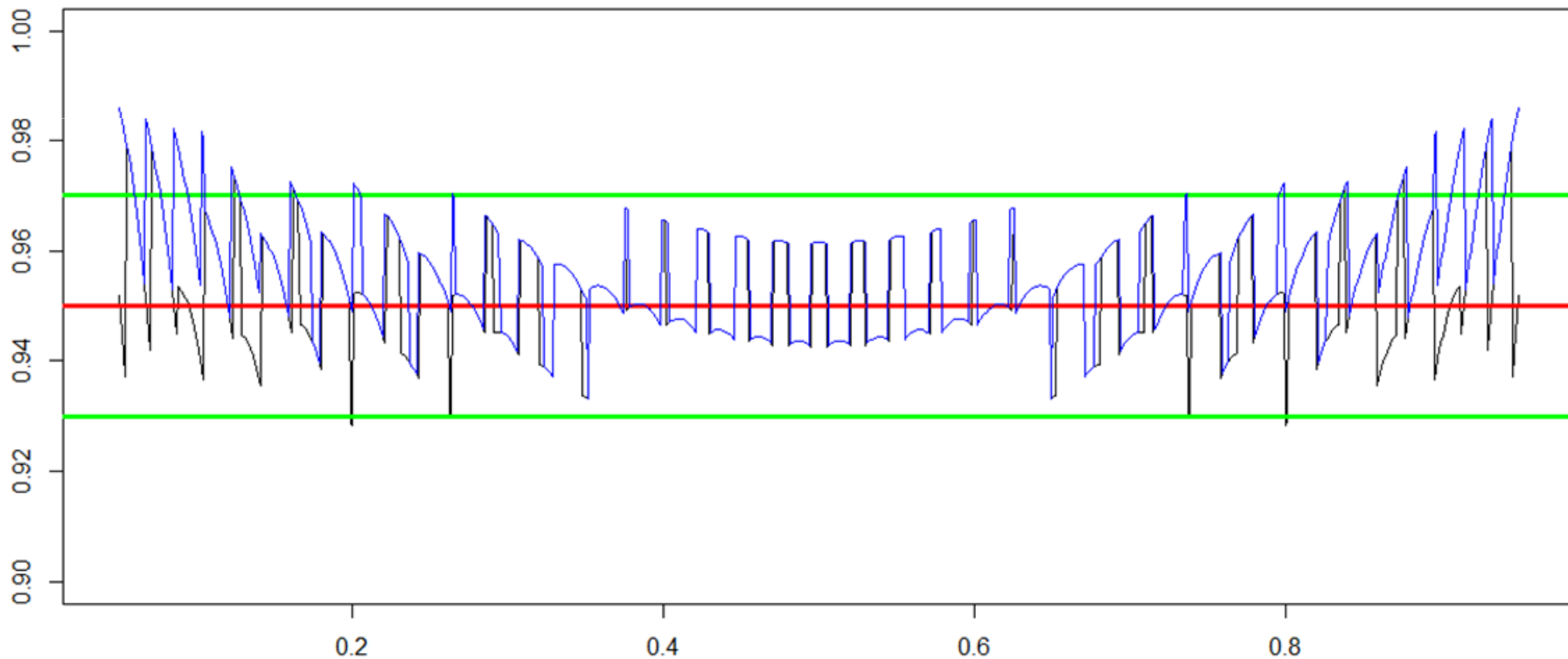
$$IC_{AC}(p, 1-\alpha) = \tilde{p} \pm u_{1-\alpha/2} \sqrt{\frac{\tilde{p}(1-\tilde{p})}{\tilde{n}}}$$

$$\alpha = 5\% \Rightarrow \tilde{T} \approx T + 2 \quad \tilde{n} \approx n + 4$$

Revient à utiliser Wald en ajoutant à l'échantillon observé 2 succès et 2 échecs

Probabilités de recouvrement de p pour Wilson (noir) et Agresti – Coull (bleu) avec $n=40$

Wilson (noir) et Agresti – Coull (bleu) : Probabilités de recouvrement pour $n=40$



Méthode X+U

- Principe :
 - On ajoute à T une variable aléatoire indépendante U de loi uniforme sur [0 1]
 - La loi de T+U a une fonction de répartition continue
 - La résolution des deux équations :

$$P(C1(p) \leq T + U \leq C2(p)) = 1 - \alpha \text{ (condition de niveau)}$$

$$E_p \left((T + U) 1_{[C1(p), C2(p)]} \right) = E_p (X + U) (1 - \alpha) \text{ (absence de biais)}$$

permet de trouver un intervalle de fluctuation [C1(p) ,C2(p)]

- L'inversion de ces intervalles de fluctuation conduit aux intervalles de confiance (cf Lehman, Testing Statistical Hypothesis)
- Les IC sont de niveau de confiance exactement égal à la valeur nominale visée (1- α) et uniformément les plus précis dans la classe des IC sans biais fondés sur X et U

Comparaison des performances des intervalles de confiance :

le développement d'Edgeworth à l'ordre 2

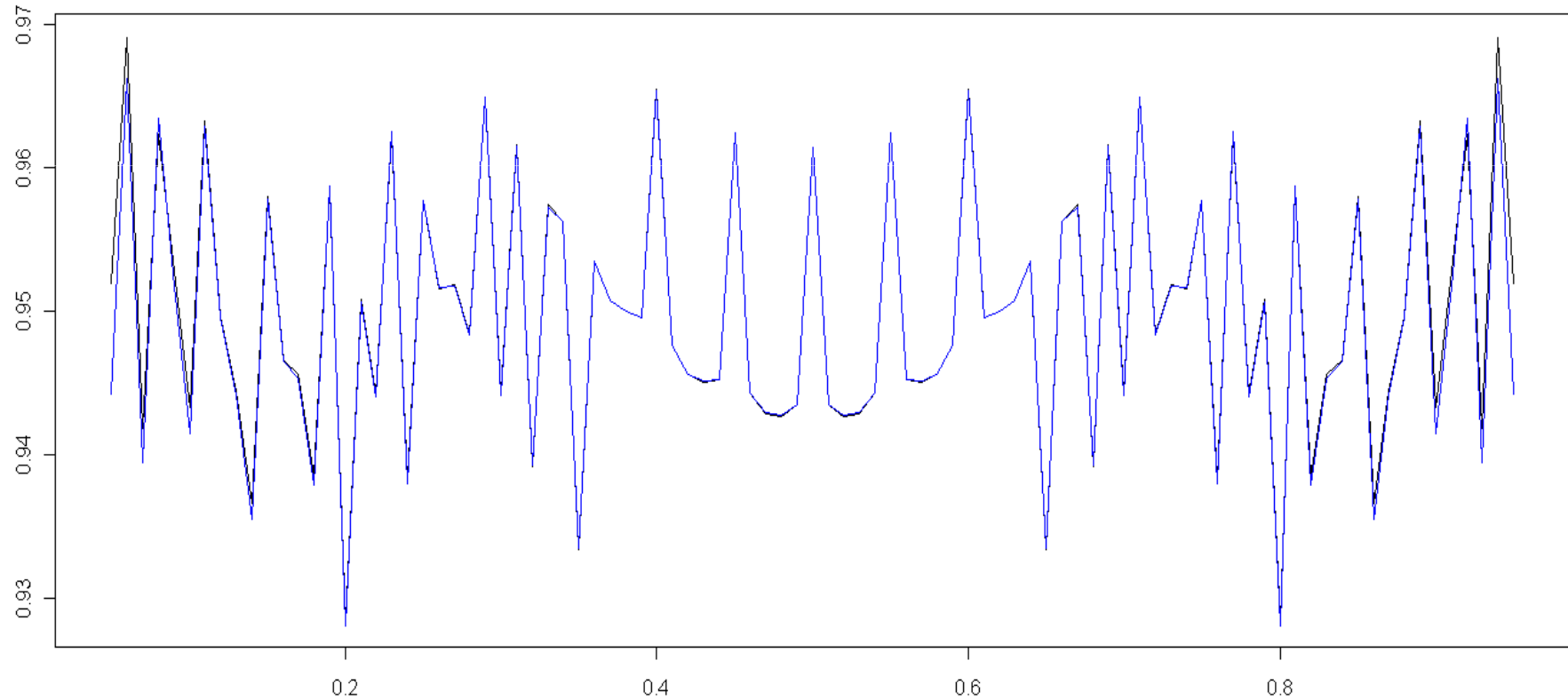
- Soit : $P(p \in IC_*) = \Pr_p(p \in [l_*(\hat{p}), u_*(\hat{p})])$
- Par un développement d'Edgeworth à l'ordre 2 de la fonction caractéristique de $Z = \frac{\sqrt{n}(\hat{p} - p)}{\sqrt{p(1-p)}}$
- On déduit une approximation analytique de $P(p \in IC_*)$

$$P(p \in IC_*) = (1 - \alpha) + \text{osc en } O(n^{-1/2}) + \text{biais en } O(n^{-1}) \\ + \text{osc en } O(n^{-1}) + O(n^{-3/2})$$

Qualité de l'approximation d'Edgeworth

Probabilité de recouvrement exacte en noir et approchée en bleu
pour Wilson $n=120$

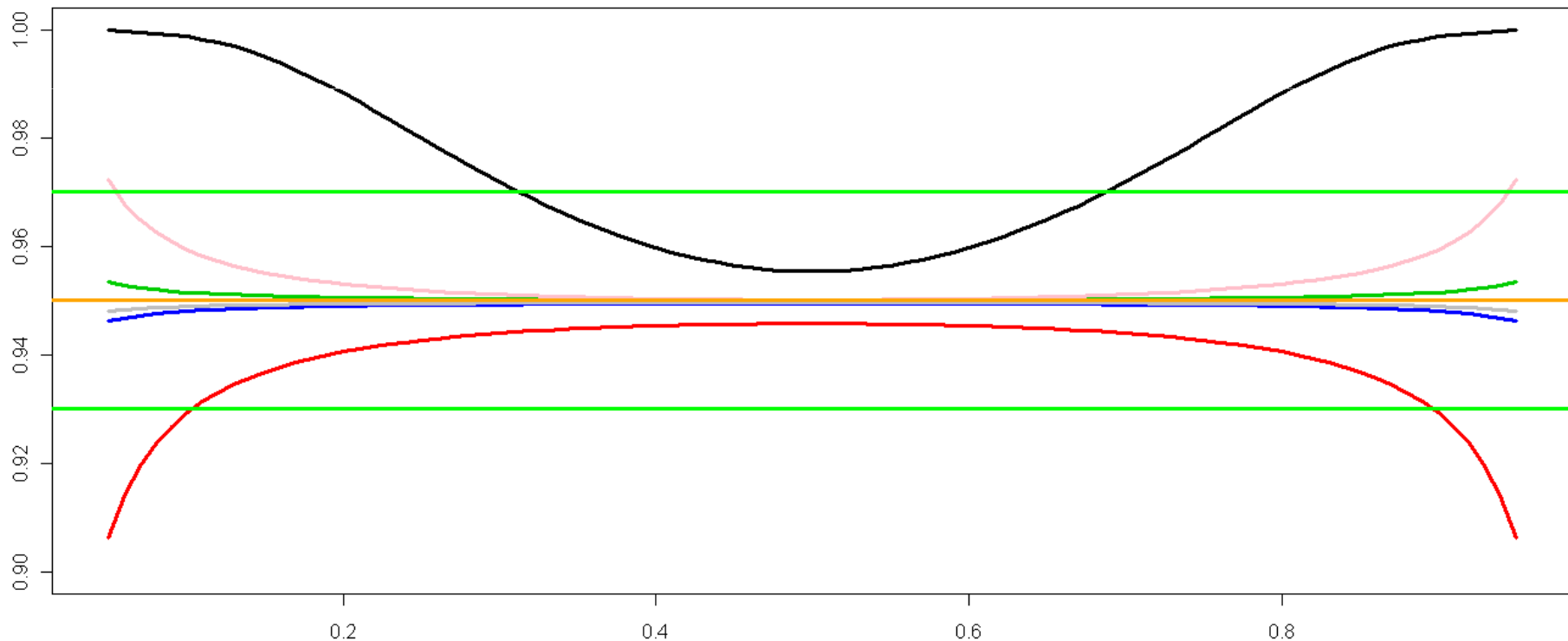
Wilson: Probabilités de recouvrement exacte (noir) et approximation d'Edgeworth (bleu) en fonction de p avec $n=120$



Probabilités de recouvrement de différentes méthodes après lissage par un développement d'Edgeworth $n=40$

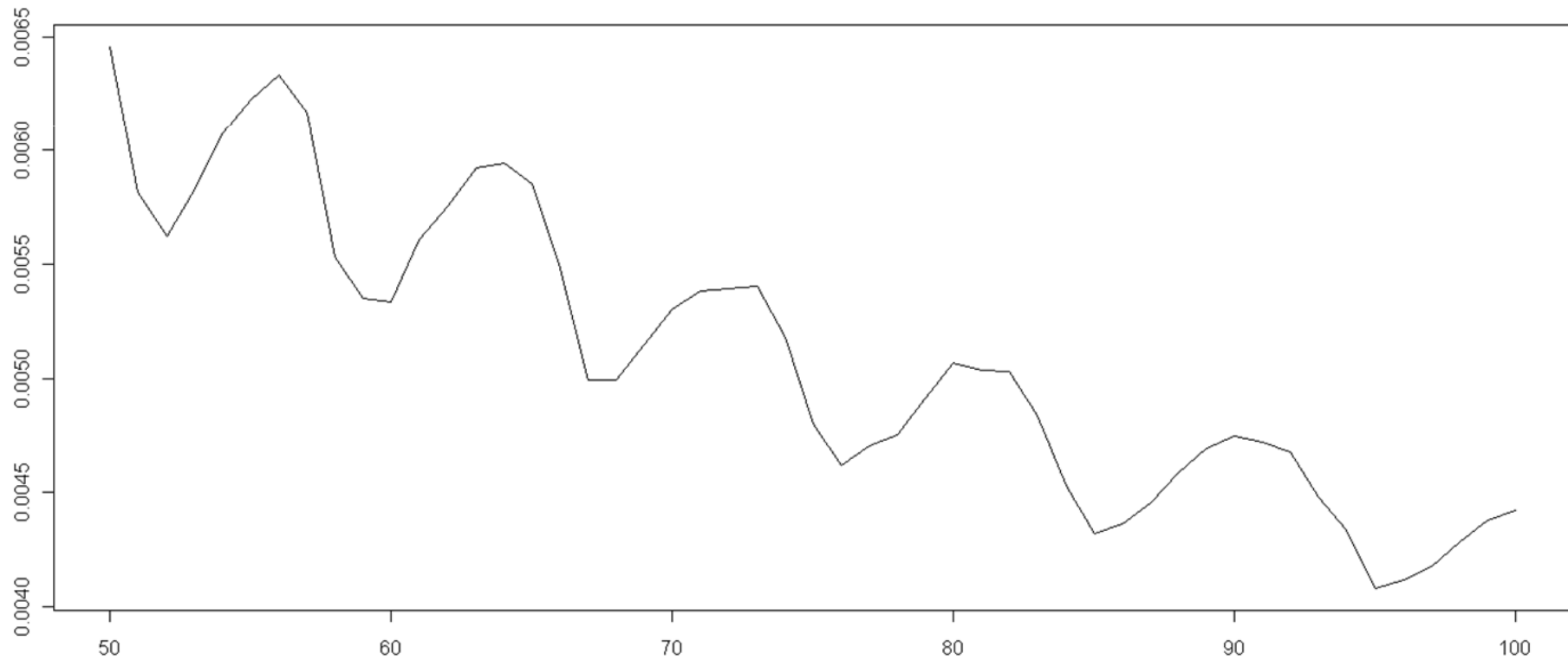
- Wald (rouge) Agresti – Coull (rose) Wilson (vert) Rapport de vraisemblance (bleu) Jeffreys (gris), X+U (orange)

approximation du biais par un développement d'Edgeworth-2 avec $n=40$: WS (rouge), RV(bleu), Jeffreys (gris), AC (rose), Wilson(vert), X+U (orange)



Amplitude moyenne des oscillations en fonction de n de 50 à 100 exemple : méthode de Wilson

oscillation en moyenne pour l'IC de Wilson en fonction de n



La précision des intervalles de confiance

- Deux critères :
 - L'espérance de la largeur de l'IC

$$L_*(n, p) = E_{n, p} [u_*(X) - l_*(X)]$$

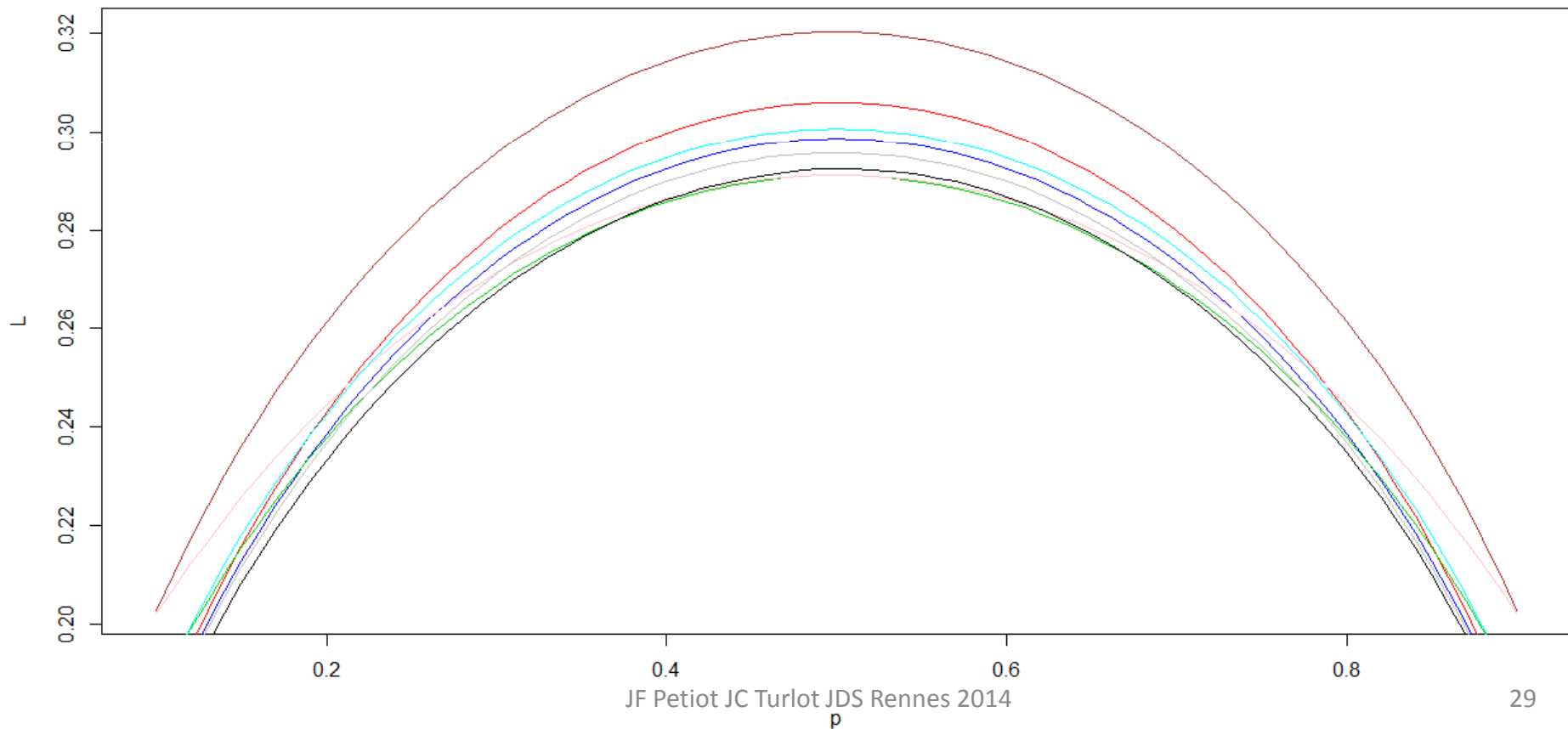
$$L_*(n, p) = \sum_{x=0}^n (u_*(x) - l_*(x)) \binom{n}{x} p^x (1-p)^{n-x}$$

- La probabilité de recouvrement des valeurs de p erronées

Espérance de la largeur de l'IC n=40

- Clopper – Pearson (brun) Wald (rouge) MID-P (bleu ciel) Rapport de vraisemblance (gris) Agresti – Coull (rose) Wilson (vert) X+U (noir)

Espérance largeur IC fonction de p avec n=40 : C-P(brun),Wald(rouge),Mid-P(bleu ciel),RV(bleu),J gris), AC(rose),Wilson(vert),X+U (noir

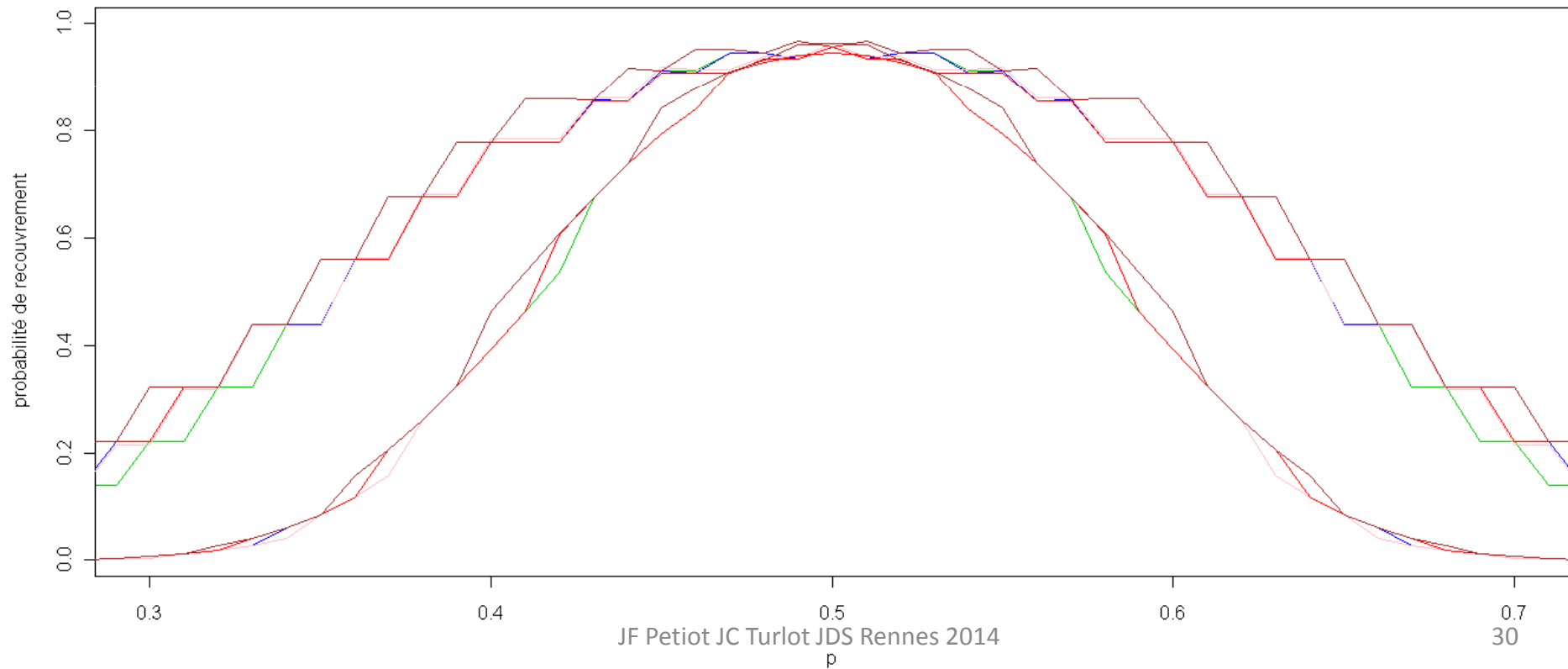


Probabilités de recouvrement de valeurs erronées de p (p vraie=0,5)

$n=40$ et $n=120$

Clopper – Pearson (brun) Wald (rouge) Rapport de vraisemblance (bleu)
Jeffreys (gris) Agresti – Coull (gris) Wilson (vert)

probabilités de recouvrement erroné $p=0.5$ avec $n=40$ et $n=120$: C-P (brun), Wald (rouge), Mid-P (bleu ciel), RV(bleu), J (gris), AC (rose), Wilson(ve.



Conclusion

| Méthode | Niveau de confiance | Largeur de l'IC | Facilité à enseigner | Facilité de mise en œuvre |
|------------------------|---------------------|-----------------|----------------------|---|
| Wald | Trop laxiste | trop large | assez facile | Très facile même sans logiciel |
| Clopper - Pearson | Trop conservatif | trop large | un peu difficile | Très facile avec un logiciel ou une table |
| MID-P | correct | correct | un peu difficile | Pas facile |
| Jeffreys | correct | correct | difficile | Facile avec un logiciel |
| Wilson | correct | précis | un peu difficile | Facile avec un logiciel |
| Agresti - Coull | correct | précis | assez facile | Très facile même sans logiciel |
| X+U | exact | précis | très difficile | Difficile |

Bibliographie

- Interval Estimation for a Binomial Proportion. L.D.BROWN, T.CAI, A. DASGUPTA. *Statistical Science*, Vol. 16 (2001), pp 101-117.
- Approximate is Better than “exact” for Interval Estimation of Binomial Proportions. A.AGRESTI, B.A.COULL. *The American Statistician*, Vol. 52 n°2 (1998) pp 119-126.
- Sample Size Requirements for the Back-of-the Envelope Binomial Confidence Interval. M.L.SAMUELS, TAI-FANG C.LU. *The American Statistician*, Vol. 46, n°3 (1992) pp228-231.
- Confidence Intervals for a Binomial Proportion and Asymptotic Expansions. L.D.BROWN, T.CAI, A. DASGUPTA. *The Annals of Statistics*, Vol. 30 (2002), pp 160-201.
- Normal Approximation and Asymptotic Expansions. BHATTACHARYA, RAO, Wiley, New-York, 1976.
- Two-sided Confidence Intervals for the single proportion; comparison of several methods. R. NEWCOMBE, *Statistics in Medicine*. 1998. 17, pp 857-872.
- Fiducial limits of the parameter of a discontinuous distribution. WL. STEVENS. *Biometrika* 1950; 37, pp 117-129.
- Teaching large-sample binomial confidence intervals. SANTNER, *Teaching Statistics*; 1998, 20, pp 20-23.
- Significance tests in discrete distributions. LANCASTER, *JASA*, 1961, 56, 223-234.
- Testing Statistical Hypothesis, E.L. LEHMANN, Wiley, 1959