



# Etude de cas, probabilités et statistique

## Rapport

Evann Semens, Logan Doceul, Tessa Paquin

2023-2024

*Cycle numérique et biologie, 2<sup>ème</sup> année*

## Sommaire

Introduction.....	4
Partie probabilités.....	4
1- Calculer la probabilité qu'un verre rempli par le serveur A contienne moins de 10cl de vin.....	4
2- Calculer la probabilité qu'un verre rempli par le serveur B contienne moins de 10cl de vin.....	4
3- Simuler un échantillon 1 de taille $n = 20, 40, 100, 200$ , verres servis par le serveur A.....	4
Résultats pour l'échantillon $n = 20$ : .....	4
Résultats pour l'échantillon $n = 40$ : .....	4
Résultats pour l'échantillon $n = 100$ : .....	5
Résultats pour l'échantillon $n = 200$ : .....	5
4- Simuler un échantillon 1 de taille $n = 20, 40, 100, 200$ , verres servis par le serveur B.....	7
Résultats pour l'échantillon $n = 20$ : .....	7
Résultats pour l'échantillon $n = 40$ : .....	7
Résultats pour l'échantillon $n = 100$ : .....	7
Résultats pour l'échantillon $n = 200$ : .....	8
Partie statistique .....	10
5- Représenter les valeurs obtenues par un graphique .....	10
Serveur A.....	10
Serveur B.....	11
6- Déterminer la proportion de verres avec moins de 10cl de vin .....	11
Serveur A.....	11
Serveur B.....	12
7- Calculer un intervalle de confiance pour la probabilité qu'un verre contienne moins de 10cl...12	12
Serveur A.....	12
Serveur B.....	12
8- Estimer $m$ et $\sigma^2$ .....	12
Serveur A.....	12
Serveur B.....	13
9- Calculer un intervalle de confiance pour $m$ .....	13
Serveur A.....	13
Serveur B.....	13
10- Faire un test de comparaison d'une moyenne à une référence pour vérifier si la moyenne est supérieure à 10cl.....	14
Serveur A.....	14
Serveur B.....	14

11-	. Faire un test de comparaison de deux moyennes afin de vérifier si les moyennes des deux serveurs sont les mêmes. ....	14
<b>Conclusion.....</b>		<b>15</b>
12-	Selon les analyses statistiques réalisées sur les données simulées, est-ce qu'il y a un problème de remplissage des verres ?.....	15
13-	Ces résultats sont-ils en accord avec les paramètres simulés ? .....	15

## Introduction

Dans cette étude de cas nous allons étudier le remplissage de verres de vin de 10cl par deux serveurs A et B. Pour ce faire nous allons utiliser le logiciel Rstudio pour modéliser les situations. Nous allons nous servir d'une graine (667) pour que toutes nos expériences aléatoires nous donnent le même résultat à chaque lecture du script.

## Partie probabilités

- 1- Calculer la probabilité qu'un verre rempli par le serveur A contienne moins de 10cl de vin

La probabilité qu'un verre rempli par le serveur A contienne moins de 10cl de vin est 0,5.

- 2- Calculer la probabilité qu'un verre rempli par le serveur B contienne moins de 10cl de vin

La probabilité qu'un verre rempli par le serveur B contienne moins de 10cl de vin est 0.2397.

- 3- Simuler un échantillon 1 de taille  $n = 20, 40, 100, 200$ , verres servis par le serveur A

Résultats pour l'échantillon  $n = 20$  :

9.571976 10.385362 10.786992 10.318636 9.764055  
10.118625 10.009651 9.599535 10.205900 9.985347  
9.681521 10.102954 10.125415 9.788633 10.711968  
9.277642 9.784978 9.489455 9.245685 10.451585

Résultats pour l'échantillon  $n = 40$  :

10.540885 10.576922 9.600791 11.126142 9.433155  
9.953818 10.089716 9.775884 10.113329 9.989159  
8.993950 10.621490 10.729424 10.569724 9.624825

9.099297 9.852353 10.677967 10.383390 10.211483  
10.106627 10.579264 10.392846 9.479546 10.258208  
10.656909 9.702307 10.397261 10.300576 9.873981  
9.225315 10.237736 9.907824 9.968545 10.037052  
10.746263 9.064622 10.040933 11.142450 10.590423

#### Résultats pour l'échantillon n= 100 :

[1] 9.770139 10.652593 10.054864 10.498552 9.683876  
[6] 10.431498 9.983686 11.087542 9.418563 10.391572  
[11] 9.216372 9.547443 10.337674 11.156904 9.843227  
[16] 11.051184 10.814263 10.508860 11.142266 9.481168  
[21] 9.295086 10.035393 10.459532 10.069486 10.237138  
[26] 9.421876 9.646635 9.915699 9.692953 9.407589  
[31] 9.713729 9.917370 10.070400 10.415854 10.608739  
[36] 9.262707 10.407958 10.118749 10.362842 9.147428  
[41] 9.213870 10.447684 9.216130 8.947690 11.484616  
[46] 9.947994 10.146963 10.403543 10.259711 9.702671  
[51] 9.706453 9.893272 9.977756 10.164300 10.408467  
[56] 10.230034 9.989290 9.505705 10.765628 9.955469

[61] 10.818421 9.782728 9.336545 10.236187 10.006622  
[66] 10.380269 10.683124 9.268384 10.151157 9.417573  
[71] 10.228648 10.720933 10.151425 9.959704 9.696476  
[76] 9.178518 9.918254 9.603889 10.254176 9.507269  
[81] 9.164901 10.153696 9.579535 10.366861 10.197311  
[86] 10.331561 9.019833 11.214700 10.309740 10.344627  
[91] 10.089633 9.830686 9.974997 9.682923 8.754467  
[96] 9.467397 10.096747 10.762095 9.586433 9.624781

#### Résultats pour l'échantillon n= 200 :

[1] 10.108374 9.907517 9.254350 10.483411 9.566991

[6] 10.971732 10.878543 9.794553 10.134523 10.371822  
[11] 9.570317 9.880044 9.973395 10.140030 9.875007  
[16] 10.151070 9.725067 9.663162 9.715084 10.163023  
[21] 9.980773 10.468253 10.468314 10.476087 9.003253  
[26] 10.078431 9.559064 9.478308 9.665970 9.156933  
[31] 9.804544 10.160142 9.663903 10.490772 9.964830  
[36] 10.241011 10.015942 9.823589 10.773970 9.196383  
[41] 10.184986 10.394191 10.232876 10.616532 10.057169  
[46] 9.946449 10.435108 11.337722 10.347430 9.879137  
[51] 10.014073 8.712435 9.952250 10.296011 10.345351  
[56] 10.546358 10.173762 8.517153 10.365979 10.469554  
[61] 10.733957 9.724588 9.539690 9.818739 9.428974  
[66] 10.342711 9.896925 10.519499 9.375383 11.336645  
[71] 10.728899 10.717202 9.348442 10.076549 10.707888  
[76] 9.287612 10.073529 9.479283 10.042591 10.094633  
[81] 9.377378 9.847576 10.203504 10.885572 9.348793  
[86] 9.273872 9.887621 9.659864 9.829540 9.208959  
[91] 11.524333 10.076726 9.586931 9.620271 10.436601  
[96] 10.089870 8.941674 10.567868 9.803996 10.257739  
[101] 9.803587 10.988070 9.467558 9.365768 10.315018  
[106] 10.420846 10.508058 9.426145 9.976667 10.490220  
[111] 9.922824 9.410024 10.924702 10.163415 10.627574  
[116] 9.728790 9.754743 9.768253 9.982195 10.435958  
[121] 10.514745 10.060484 10.063964 10.262488 9.845827  
[126] 10.029552 9.809796 9.167760 10.342539 10.916189  
[131] 9.356916 9.670421 9.957151 10.240003 10.277275  
[136] 10.709927 9.892202 10.454064 10.198498 10.215007  
[141] 10.108712 9.283242 10.041562 10.970125 10.069691  
[146] 10.667301 9.536739 10.000032 9.688063 9.453430  
[151] 9.395408 10.265826 9.622853 10.600731 11.136866  
[156] 10.271217 10.393316 9.771304 10.130900 9.789675  
[161] 10.659966 9.835917 10.303983 10.412030 10.934687

[166] 9.228575 10.402378 10.275532 10.300585 9.664319  
 [171] 9.086550 11.145777 10.104281 10.232451 10.310059  
 [176] 9.266474 10.367754 9.711884 10.346794 9.416894  
 [181] 10.004572 10.151066 9.623854 9.922653 9.648175  
 [186] 10.186948 9.997375 9.689451 10.510513 10.228254  
 [191] 10.169124 10.135732 8.801634 10.370614 10.599140  
 [196] 9.291646 9.970746 9.974931 9.732416 10.020064

#### 4- Simuler un échantillon 1 de taille $n = 20, 40, 100, 200$ , verres servis par le serveur B

Résultats pour l'échantillon  $n= 20$  :

[1] 10.291182 10.686948 10.937744 10.659517 11.126612  
 [6] 10.870294 11.782778 10.178929 9.967034 10.682836  
 [11] 10.473017 10.611760 9.796247 10.421232 10.096871  
 [16] 11.167013 10.161800 10.163918 10.354354 10.184207

Résultats pour l'échantillon  $n= 40$  :

[1] 10.526169 9.589422 10.406844 11.316413 11.158793  
 [6] 10.438942 10.384043 10.108652 10.678148 11.718770  
 [11] 9.928411 9.830945 10.660139 9.645493 10.814474  
 [16] 9.930937 9.351592 9.603237 11.370335 10.307204  
 [21] 11.231184 10.634956 11.112593 10.600263 10.285794  
 [26] 10.815353 10.575481 10.904488 10.137279 10.859821  
 [31] 10.396450 9.714976 10.680605 10.047696 10.540793  
 [36] 10.165226 9.668842 10.659902 10.393626 11.285372

Résultats pour l'échantillon  $n= 100$  :

[1] 10.327539 11.549862 11.046948 10.674879 10.548719  
 [6] 10.055876 11.722667 10.478982 10.761857 9.690899  
 [11] 9.664263 10.490159 10.326390 9.687629 11.065609

[16] 10.209417 10.805424 11.538983 9.958953 10.844087  
 [21] 10.142790 11.213084 10.630576 10.772176 10.468214  
 [26] 10.343116 10.034477 9.353525 10.948399 10.761695  
 [31] 10.599150 10.975183 10.475473 9.575975 11.020022  
 [36] 10.424363 10.124291 10.175192 10.346994 10.135440  
 [41] 11.827098 10.846200 9.664568 10.649433 11.633833  
 [46] 10.422889 10.886639 10.413611 10.768117 10.788253  
 [51] 10.458921 10.703339 10.530603 10.943305 11.135079  
 [56] 9.842304 10.398772 9.534615 9.908873 10.221326  
 [61] 10.719942 10.030863 10.430443 10.764144 11.046095  
 [66] 10.198827 10.806474 10.043214 10.330439 11.208424  
 [71] 10.610532 11.293060 10.713810 10.216171 9.785397  
 [76] 11.010183 11.139775 10.650606 9.815547 10.551212  
 [81] 10.176944 11.045594 9.607335 9.441244 10.717973  
 [86] 10.363346 10.810107 10.843176 11.152387 10.171567  
 [91] 10.112007 9.883377 9.994028 10.919341 10.209609  
 [96] 10.346961 11.747792 10.779371 10.410167 10.119873

#### Résultats pour l'échantillon n= 200 :

[1] 10.961399 10.364551 10.891130 10.246711 10.704306  
 [6] 10.232761 10.703695 10.541299 10.652864 10.346791  
 [11] 10.296507 10.480218 9.446692 9.610156 9.728541  
 [16] 10.296027 10.023079 9.728559 10.697716 10.747529  
 [21] 10.360880 10.443762 10.031825 9.535351 10.060929  
 [26] 10.372678 10.618485 10.595931 10.348161 10.423177  
 [31] 10.800159 10.280007 10.067079 8.900433 11.234069  
 [36] 11.168173 10.421211 10.162226 11.461654 10.404577  
 [41] 10.207018 10.583722 10.583227 10.469326 10.909621  
 [46] 11.645189 10.494939 11.189912 10.614919 10.913767  
 [51] 10.116564 10.344921 10.817353 10.505207 11.167487  
 [56] 9.522402 10.829020 9.658316 10.715989 9.690000  
 [61] 9.707156 11.523651 10.343847 10.371582 10.677157

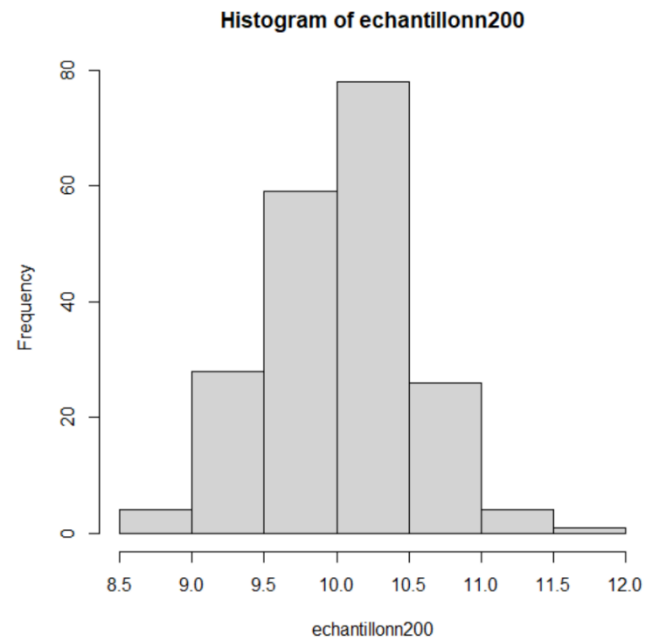
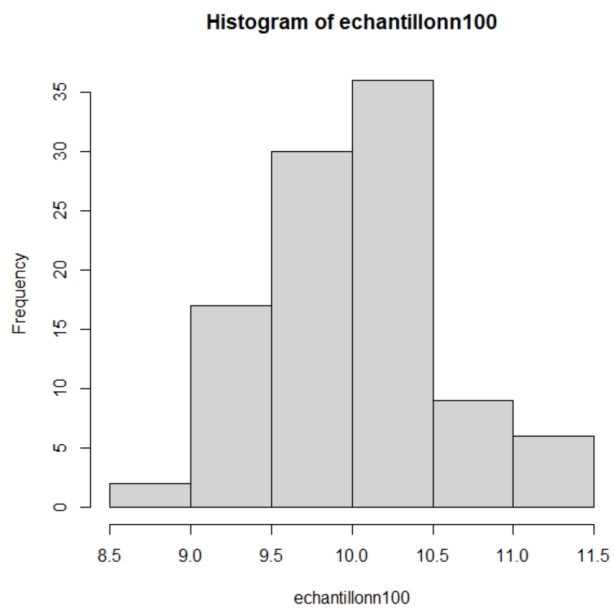
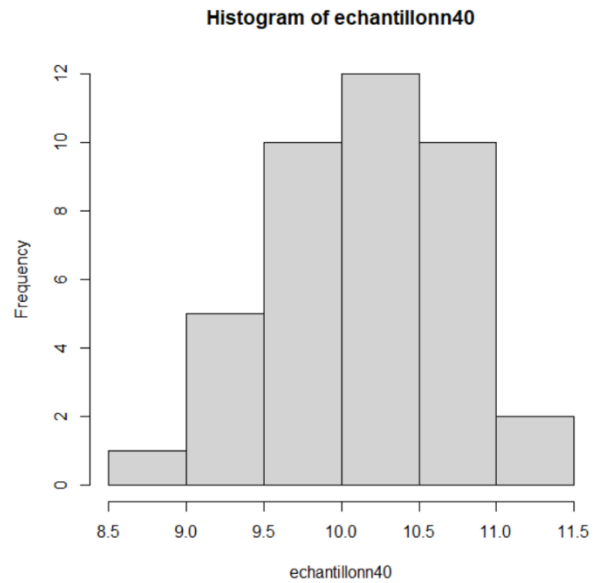
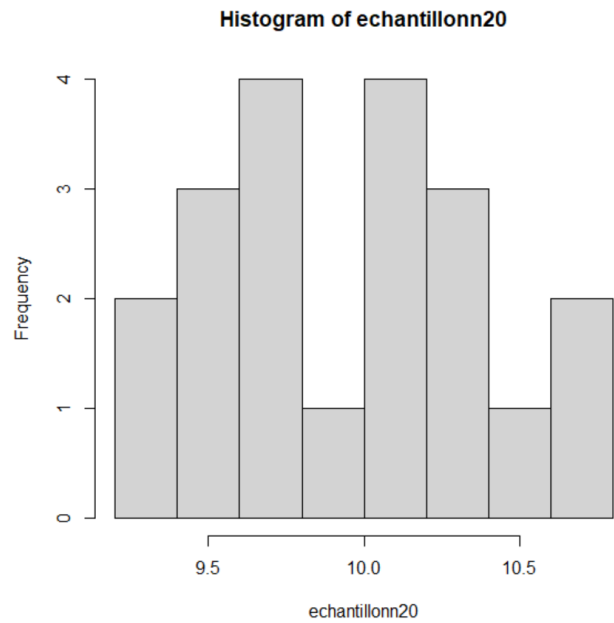


[66] 11.236001 10.327364 10.384038 9.884353 10.394273  
[71] 9.978484 10.476218 11.046212 9.963144 9.474791  
[76] 10.720811 10.516846 10.116785 10.553181 11.004576  
[81] 10.547375 10.642405 10.844913 10.203680 10.534388  
[86] 9.798751 10.799063 10.703671 11.778866 10.280627  
[91] 10.742256 10.504341 10.587844 9.877029 10.925639  
[96] 10.549265 10.755371 10.646839 10.374558 9.949079  
[101] 10.309614 11.045836 11.270711 11.377905 10.340369  
[106] 10.905343 10.487307 10.043243 10.731355 10.308583  
[111] 10.254837 10.588484 11.389829 10.641160 10.261514  
[116] 10.581578 9.222763 10.658910 10.974912 10.850279  
[121] 10.848320 10.517351 10.342715 11.385024 10.198041  
[126] 10.193799 10.682701 9.948722 10.085014 10.538547  
[131] 11.314883 10.143539 10.139252 10.293999 9.719252  
[136] 10.567530 10.905253 10.525679 10.819463 10.291967  
[141] 10.367902 10.569152 9.196241 10.476060 11.421298  
[146] 10.335805 9.653685 10.448808 10.080632 10.723005  
[151] 9.538127 10.728333 10.790198 10.471965 10.973298  
[156] 10.425513 10.941670 10.665557 10.345193 10.222461  
[161] 10.669385 11.188096 10.150167 9.839288 10.214834  
[166] 10.967482 10.230071 10.101302 10.390283 10.989035  
[171] 10.955479 10.357837 10.761413 9.839866 10.073039  
[176] 10.701132 11.325031 11.236671 10.451619 11.220215  
[181] 10.395799 10.337598 10.524345 10.213263 10.092954  
[186] 10.472416 10.669497 10.884223 11.352285 10.991121  
[191] 9.605184 10.073718 10.744277 10.020312 10.385991  
[196] 10.585451 10.081334 10.273051 10.277970 10.364664

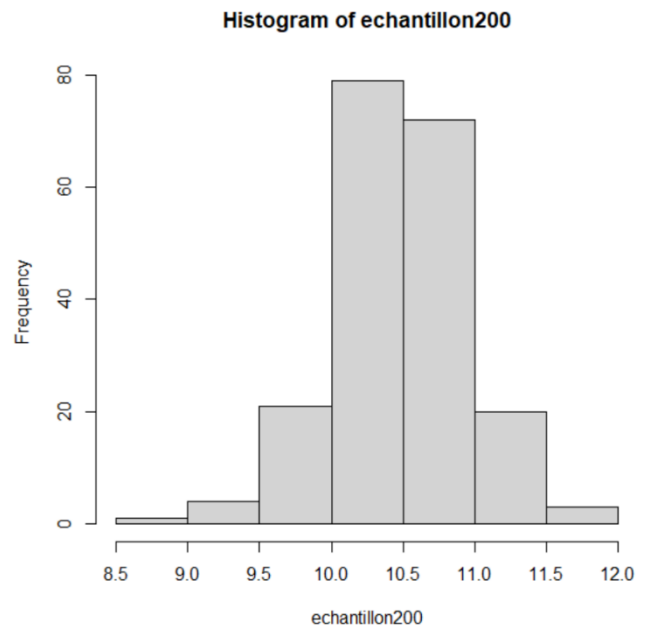
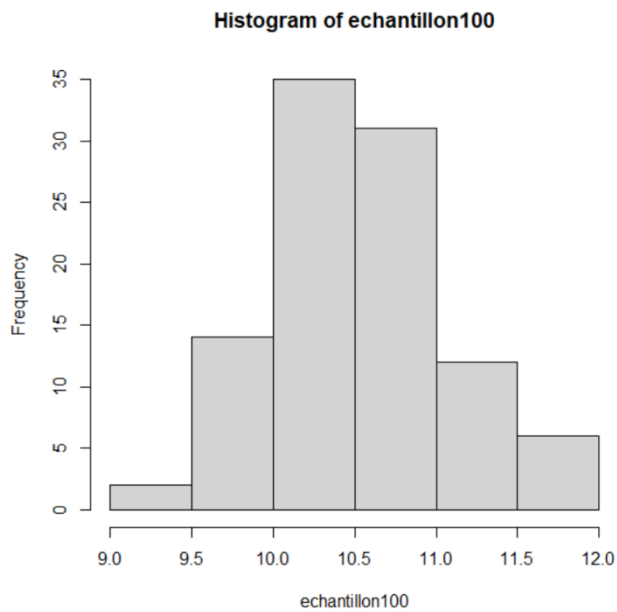
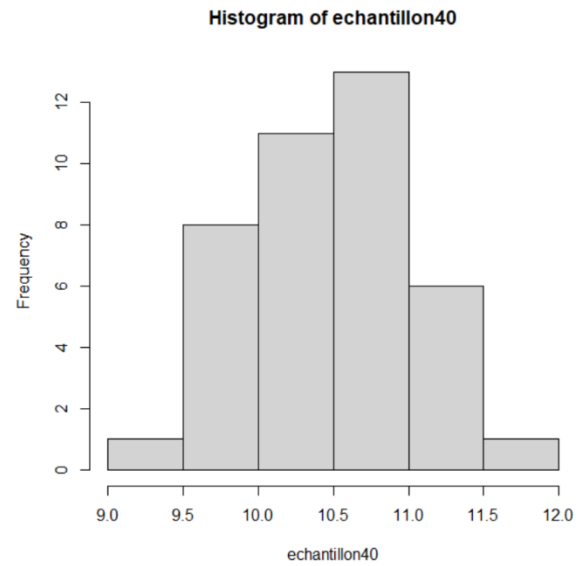
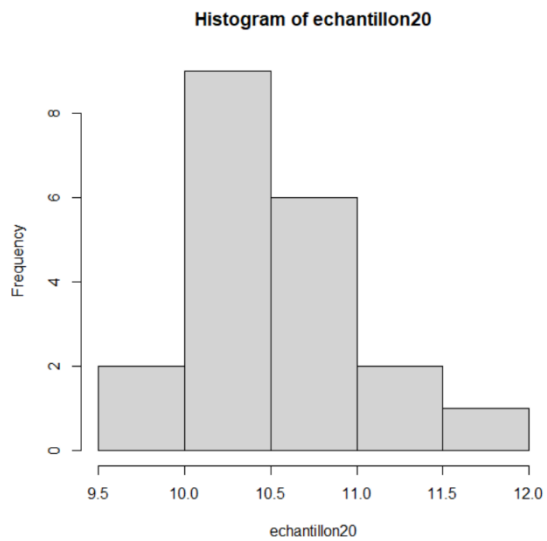
## Partie statistique

### 5- Représenter les valeurs obtenues par un graphique

Serveur A



## Serveur B



## 6- Déterminer la proportion de verres avec moins de 10cl de vin

### Serveur A

Soit  $P$  la proportion de verres avec moins de 10 cl de vin,

$N = 20$ ,  $P = 50\%$

$N = 40$ ,  $P = 40\%$

$N = 100$ ,  $P = 49\%$

$N = 200$ ,  $P = 45\%$

### Serveur B

Soit P la proportion de verres avec moins de 10 cl de vin,

N= 20, **P= 10%**

N= 40, **P = 22%**

N=100, **P= 16%**

N = 200, **P= 13%**

## 7- Calculer un intervalle de confiance pour la probabilité qu'un verre contienne moins de 10cl

### Serveur A

N= 20

**[9,8071 ; 10,1928]**

N = 40

**[9,8668 ; 10,1331]**

N=100

**[9,91699 ; 10,08301]**

N= 200

**[9,9416 ; 10,0584]**

### Serveur B

N= 20

**[9,8071 ; 10.71557]**

N = 40

**[9,8668 ; 10.61145]**

N=100

**[9,91699 ; 10.60626]**

N= 200

**[9,9416 ; 10.53212]**

## 8- Estimer m et $\sigma^2$

### Serveur A

N=20, **m=9,97,  $\sigma^2 = 0.1888793$**

**N=40, m= 10.11681,  $\sigma^2 = 0.2861505$**

**N=100, m= 10.015,  $\sigma^2 = 0.3029774$**

**N= 200, m= 10.03512,  $\sigma^2 = 0.2623093$**

#### Serveur B

**N=20, m=10.53071,  $\sigma^2 = 0.2285778$**

**N=40, m= 10.46199,  $\sigma^2 = 0.3147588$**

**N=100, m=10.51764,  $\sigma^2 = 0.2848167$**

**N= 200, m= 10.47602,  $\sigma^2 = 0.2304804$**

### 9- Calculer un intervalle de confiance pour m

#### Serveur A

N=20 :

[9.766896 ; 10.173696]

N=40 :

[9.945731 ; 10.287889]

N=100 :

[9.905781 ; 10.124217]

N=200 :

[9.963702 ; 10.106532]

#### Serveur B

N=20 :

[10.30696 ; 10.75447]

N=40 :

[10.28256 ; 10.64142]

N=100 :

[10.41175 ; 10.6235]

N=200 :

[10.40908 ; 10.54297]

10-Faire un test de comparaison d'une moyenne à une référence pour vérifier si la moyenne est supérieure à 10cl.

Serveur A

On prend un risque alpha de 5%, l'échantillon étant grand on peut faire le test de comparaison d'une moyenne à une référence.

Hypothèses :  $H_0 : \mu_1 = \mu_2$

$H_1 : \mu_1$  différent de  $\mu_2$

Décision : La p-value étant inférieure au risque alpha, on rejette  $H_0$  et donc les moyennes sont significativement différentes à 10.

p-value = 0,3 < 0,5

N=20 :

La moyenne est inférieure à 10 selon le test.

N=40 :

La moyenne est supérieure à 10 selon le test.

N=100 :

La moyenne est supérieure à 10 selon le test.

N=200 :

La moyenne est supérieure à 10 selon le test.

Serveur B

La moyenne est supérieure à 10 selon le test.

11-. Faire un test de comparaison de deux moyennes afin de vérifier si les moyennes des deux serveurs sont les mêmes.

On prend un risque alpha de 5%.

Conditions d'application : l'échantillon est grand, on peut faire un test de comparaison de deux moyennes.

Hypothèses :

$H_0 : \mu_1 = \mu_2$

$H_1 : \mu_1$  différent de  $\mu_2$

Décision : On trouve une p-value inférieure au risque alpha donc on rejette  $H_0$ , les moyennes sont significativement différentes. ( $0.08756 < 0,5$ )

$N=20$ , On obtient du test de comparaison de deux moyennes que les moyennes des deux serveurs sont différentes. (9.970296 et 10.530715)

$N=40$ , On obtient du test de comparaison de deux moyennes que les moyennes des deux serveurs sont différentes. (10.11681 et 10.46199)

$N=100$ , On obtient du test de comparaison de deux moyennes que les moyennes des deux serveurs sont différentes. (10.01500 et 10.51764)

$N=200$ , On obtient du test de comparaison de deux moyennes que les moyennes des deux serveurs sont différentes. (10.03512 et 10.47602)

## Conclusion

12-Selon les analyses statistiques réalisées sur les données simulées, est-ce qu'il y a un problème de remplissage des verres ?

Selon les analyses statistiques on peut dire qu'il y a un problème de remplissage des verres entre les deux serveurs. En effet, en comparant les deux moyennes de remplissage des verres grâce à un test de comparaison de deux moyennes que la moyenne de remplissage de verre du serveur A est inférieure à celle du serveur B.

13-Ces résultats sont-ils en accord avec les paramètres simulés ?

Ces résultats sont en accord avec les paramètres simulés, en effet, on a pu voir que la moyenne de remplissage du serveur A était inférieure à 10 alors que celle du serveur B

est supérieure. De plus la proportion de verres qui contiennent moins de 10cl de vin est bien moins importante chez le serveur B que chez le serveur A.