

Démographie

Calcul d'un quantile et de l'espérance de vie à partir d'une table de mortalité

■ Résumé

A partir d'une table de mortalité, pour une personne d'un âge donné x , comment calculer

- un quantile, par exemple l'âge y tel que la probabilité d'atteindre l'âge y est de 80 %;
- l'espérance de vie ?

La méthode de calcul fait appel à la représentation des données par un histogramme, plus précisément

- les âges sont subdivisés en intervalles d'une année: la variable "**borne**" définit des bornes des classes;
- pour chaque classe, la probabilité de décès "**fd**" est calculée;
les *Tables de mortalité* donnant des probabilités conditionnelles, on peut représenter la situation par un arbre;
chaque "**fd**" est un produit de probabilités conditionnelles;
- les fréquences cumulées "**ffd**" représentent des valeurs de la fonction de distribution aux "**borne**" correspondantes;
- un quantile se calcule avec la fonction réciproque de la fonction de distribution, c'est-à-dire en intervertissant les rôles de "**borne**" et de "**ffd**";
pour obtenir une valeur intermédiaire, on interpole linéairement dans la table.

■ Liens hypertextes

vers le calculateur en ligne

Espérance de vie d'un homme ou d'une femme en fonction de l'âge, avec quintiles de mortalité et histogramme
<http://www.deleze.name/marcel/culture/probabilites/population/index.html>

vers la page mère

<http://www.deleze.name/marcel/sec2/applmaths/index.html>

■ Sources

Cours: *Statistique descriptive à une variable continue*

http://www.deleze.name/marcel/sec2/applmaths/csud/statistique_1/2-stat_I.pdf

Tables de mortalité de l'*Office fédéral de la Statistique* des hommes en 2009.

■ Input

```
sexe = "homme" ;
```

```
age = 39.25 ;
```

■ Table de mortalité et probabilité conditionnelle de décès "pcd"

tpd = table des probabilités de décès de l'âge x à l'âge $x+1$;

tpd[x] = probabilité qu'un homme soit décédé à l'âge $x+1$ sachant qu'il est vivant à l'âge x ,

pour $x = 0, 1, 2, \dots, 99$

Exemple: hommes en 2009

```

tpd = {0.004811, 0.000200, 0.000108, 0.000093, 0.000084, 0.000079, 0.000076, 0.000077,
0.000079, 0.000083, 0.000088, 0.000095, 0.000103, 0.000112, 0.000128, 0.000159,
0.000222, 0.000328, 0.000477, 0.000633, 0.000727, 0.000740, 0.000701,
0.000649, 0.000609, 0.000583, 0.000570, 0.000566, 0.000570, 0.000582,
0.000600, 0.000625, 0.000654, 0.000688, 0.000726, 0.000768, 0.000815,
0.000868, 0.000927, 0.000995, 0.001072, 0.001160, 0.001261, 0.001379,
0.001515, 0.001672, 0.001852, 0.002057, 0.002288, 0.002547, 0.002835,
0.003152, 0.003498, 0.003870, 0.004267, 0.004688, 0.005136, 0.005614,
0.006126, 0.006676, 0.007269, 0.007913, 0.008616, 0.009388, 0.010243,
0.011191, 0.012248, 0.013430, 0.014756, 0.016249, 0.017938, 0.019854,
0.022037, 0.024535, 0.027401, 0.030687, 0.034447, 0.038738, 0.043624,
0.049172, 0.055451, 0.062531, 0.070484, 0.079374, 0.089262, 0.100185,
0.112164, 0.125189, 0.139220, 0.154175, 0.169926, 0.186290, 0.203031,
0.219854, 0.235959, 0.248497, 0.253754, 0.248281, 0.230005, 0.199353};

```

```
ltpd = Length[tpd]
```

```
100
```

■ Bornes des intervalles des classes

Seule la partie des tables postérieure à l'âge de la personne est exploitée.

En fin de liste, une classe est rajoutée pour les personnes qui vivent au-delà du dernier âge tabulé.

```
zero = Floor[age]
```

```
39
```

```
un = zero + 1
```

```
40
```

```
Clear[borne, i]; borne[zero] = age;
```

```
Do[borne[i] = i, {i, un, ltpd}];
```

```
borne[ltpd + 1] = 105;
```

```
? borne
```

```
Global`borne
```

```
borne[39] = 39.25
```

```
borne[40] = 40
```

```
borne[41] = 41
```

```
borne[42] = 42
```

```
borne[43] = 43
```

```
borne[44] = 44
```

```
borne[45] = 45
```

```
borne[46] = 46
```

```
borne[47] = 47
```

```
borne[48] = 48
```

```
borne[49] = 49
```

```
borne[50] = 50
```

```
borne[51] = 51
```

borne[52] = 52
borne[53] = 53
borne[54] = 54
borne[55] = 55
borne[56] = 56
borne[57] = 57
borne[58] = 58
borne[59] = 59
borne[60] = 60
borne[61] = 61
borne[62] = 62
borne[63] = 63
borne[64] = 64
borne[65] = 65
borne[66] = 66
borne[67] = 67
borne[68] = 68
borne[69] = 69
borne[70] = 70
borne[71] = 71
borne[72] = 72
borne[73] = 73
borne[74] = 74
borne[75] = 75
borne[76] = 76
borne[77] = 77
borne[78] = 78
borne[79] = 79
borne[80] = 80
borne[81] = 81
borne[82] = 82
borne[83] = 83
borne[84] = 84
borne[85] = 85
borne[86] = 86
borne[87] = 87
borne[88] = 88
borne[89] = 89
borne[90] = 90
borne[91] = 91
borne[92] = 92
borne[93] = 93
borne[94] = 94
borne[95] = 95
borne[96] = 96

```

borne[97] = 97
borne[98] = 98
borne[99] = 99
borne[100] = 100
borne[101] = 105

```

■ Probabilité conditionnelle de décès "pcd"

Seule la partie des tables postérieure à l'âge de la personne est exploitée. Pour un âge qui n'est pas entier, la fréquence de la classe correspondante est fractionnée.

pcd = probabilité conditionnelle de décès, comptabilisée à partir de l'âge donné;

pcd[i] = probabilité que la personne soit vivante à l'âge **borne[i+1]** sachant qu'elle est vivante à l'âge **borne[i]**

```

Clear[pcd, i]; pcd[zero] = tpd[[zero + 1]] * (un - age);
Do[pcd[i] = tpd[[i + 1]], {i, un, 1tpd - 1}];
? pcd

```

```
Global`pcd
```

```

pcd[39] = 0.00074625
pcd[40] = 0.001072
pcd[41] = 0.00116
pcd[42] = 0.001261
pcd[43] = 0.001379
pcd[44] = 0.001515
pcd[45] = 0.001672
pcd[46] = 0.001852
pcd[47] = 0.002057
pcd[48] = 0.002288
pcd[49] = 0.002547
pcd[50] = 0.002835
pcd[51] = 0.003152
pcd[52] = 0.003498
pcd[53] = 0.00387
pcd[54] = 0.004267
pcd[55] = 0.004688
pcd[56] = 0.005136
pcd[57] = 0.005614
pcd[58] = 0.006126
pcd[59] = 0.006676
pcd[60] = 0.007269
pcd[61] = 0.007913
pcd[62] = 0.008616
pcd[63] = 0.009388
pcd[64] = 0.010243
pcd[65] = 0.011191

```

```
pcd[66] = 0.012248
pcd[67] = 0.01343
pcd[68] = 0.014756
pcd[69] = 0.016249
pcd[70] = 0.017938
pcd[71] = 0.019854
pcd[72] = 0.022037
pcd[73] = 0.024535
pcd[74] = 0.027401
pcd[75] = 0.030687
pcd[76] = 0.034447
pcd[77] = 0.038738
pcd[78] = 0.043624
pcd[79] = 0.049172
pcd[80] = 0.055451
pcd[81] = 0.062531
pcd[82] = 0.070484
pcd[83] = 0.079374
pcd[84] = 0.089262
pcd[85] = 0.100185
pcd[86] = 0.112164
pcd[87] = 0.125189
pcd[88] = 0.13922
pcd[89] = 0.154175
pcd[90] = 0.169926
pcd[91] = 0.18629
pcd[92] = 0.203031
pcd[93] = 0.219854
pcd[94] = 0.235959
pcd[95] = 0.248497
pcd[96] = 0.253754
pcd[97] = 0.248281
pcd[98] = 0.230005
pcd[99] = 0.199353
```

■ Probabilité de survie "ps"

La probabilité que la personne soit vivante à l'âge **borne[i]** est égale au produit des probabilités suivantes

* (probabilité que la personne soit vivante à l'âge **borne[zero+1]**) = **(1 - pcd[zero])**

* (probabilité que la personne soit vivante à l'âge **borne[zero+2]** sachant qu'elle était vivante à l'âge **borne[zero+1]**)
= **(1 - pcd[zero+1])**

* ...

* (probabilité que la personne soit vivante à l'âge **borne[i]** sachant qu'elle était vivante à l'âge **borne[i-1]**)
= **(1 - pcd[i-1])**

$$ps[i] = (1 - pcd[zero]) * (1 - pcd[zero + 1]) * \dots * (1 - pcd[i - 1])$$

Remarque : le nombre de survivants finit par être nul.

```
Clear[ps, i]; ps[zero] = 1;  
Do[ps[i] = ps[i - 1] * (1 - pcd[i - 1]), {i, un, ltpd}];  
ps[ltpd + 1] = 0;  
? ps
```

```
Global`ps
```

```
ps[39] = 1  
ps[40] = 0.999254  
ps[41] = 0.998183  
ps[42] = 0.997025  
ps[43] = 0.995767  
ps[44] = 0.994394  
ps[45] = 0.992888  
ps[46] = 0.991228  
ps[47] = 0.989392  
ps[48] = 0.987357  
ps[49] = 0.985098  
ps[50] = 0.982589  
ps[51] = 0.979803  
ps[52] = 0.976715  
ps[53] = 0.973298  
ps[54] = 0.969531  
ps[55] = 0.965394  
ps[56] = 0.960869  
ps[57] = 0.955934  
ps[58] = 0.950567  
ps[59] = 0.944744  
ps[60] = 0.938437  
ps[61] = 0.931615  
ps[62] = 0.924243  
ps[63] = 0.91628  
ps[64] = 0.907678  
ps[65] = 0.898381  
ps[66] = 0.888327  
ps[67] = 0.877447  
ps[68] = 0.865663  
ps[69] = 0.852889  
ps[70] = 0.83903  
ps[71] = 0.82398  
ps[72] = 0.80762  
ps[73] = 0.789823
```

```
ps[74] = 0.770445
ps[75] = 0.749334
ps[76] = 0.726339
ps[77] = 0.701319
ps[78] = 0.674151
ps[79] = 0.644742
ps[80] = 0.613039
ps[81] = 0.579045
ps[82] = 0.542837
ps[83] = 0.504575
ps[84] = 0.464525
ps[85] = 0.423061
ps[86] = 0.380676
ps[87] = 0.337978
ps[88] = 0.295667
ps[89] = 0.254504
ps[90] = 0.215266
ps[91] = 0.178687
ps[92] = 0.145399
ps[93] = 0.115879
ps[94] = 0.0904023
ps[95] = 0.0690711
ps[96] = 0.0519071
ps[97] = 0.0387355
ps[98] = 0.0291182
ps[99] = 0.0224209
ps[100] = 0.0179512
ps[101] = 0
```

■ Probabilité de décès, ou fréquence cumulée des classes

$\text{ffd}[i]$ = (probabilité que la personne soit décédée à l'âge $\text{borne}[i]$) = $(1 - \text{ps}[i])$

```
Clear[ffd]; Do[ffd[i] = 1 - ps[i], {i, zero, ltpd + 1}]; ? ffd
```

```
Global`ffd
```

```
ffd[39] = 0
ffd[40] = 0.00074625
ffd[41] = 0.00181745
ffd[42] = 0.00297534
ffd[43] = 0.00423259
ffd[44] = 0.00560575
ffd[45] = 0.00711226
ffd[46] = 0.00877237
ffd[47] = 0.0106081
```

ffd[48] = 0.0126433
ffd[49] = 0.0149024
ffd[50] = 0.0174114
ffd[51] = 0.0201971
ffd[52] = 0.0232854
ffd[53] = 0.0267019
ffd[54] = 0.0304686
ffd[55] = 0.0346056
ffd[56] = 0.0391314
ffd[57] = 0.0440664
ffd[58] = 0.049433
ffd[59] = 0.0552562
ffd[60] = 0.0615633
ffd[61] = 0.0683848
ffd[62] = 0.0757566
ffd[63] = 0.0837199
ffd[64] = 0.092322
ffd[65] = 0.101619
ffd[66] = 0.111673
ffd[67] = 0.122553
ffd[68] = 0.134337
ffd[69] = 0.147111
ffd[70] = 0.16097
ffd[71] = 0.17602
ffd[72] = 0.19238
ffd[73] = 0.210177
ffd[74] = 0.229555
ffd[75] = 0.250666
ffd[76] = 0.273661
ffd[77] = 0.298681
ffd[78] = 0.325849
ffd[79] = 0.355258
ffd[80] = 0.386961
ffd[81] = 0.420955
ffd[82] = 0.457163
ffd[83] = 0.495425
ffd[84] = 0.535475
ffd[85] = 0.576939
ffd[86] = 0.619324
ffd[87] = 0.662022
ffd[88] = 0.704333
ffd[89] = 0.745496
ffd[90] = 0.784734
ffd[91] = 0.821313


```

ffd[92] = 0.854601
ffd[93] = 0.884121
ffd[94] = 0.909598
ffd[95] = 0.930929
ffd[96] = 0.948093
ffd[97] = 0.961265
ffd[98] = 0.970882
ffd[99] = 0.977579
ffd[100] = 0.982049
ffd[101] = 1

```

■ Fréquences des classes

fd[i] = fréquence de décès de la classe d'âge de **borne[i]** à **borne[i+1]**

Les *Tables de mortalité* donnant des probabilités conditionnelles, on peut représenter la situation par un arbre; chaque **fd[i]** est un produit de probabilités conditionnelles.

```

Clear[fd, i];
Do[fd[i] = ps[i] * pcd[i], {i, zero, ltpd}]

```

Les fréquences des classes sont aussi les différences successives de la fréquence cumulée

$$fd[i] = ffd[i + 1] - ffd[i]$$

ce qui nous permet de calculer la fréquence de la dernière classe

```
fd[ltpd] = ffd[ltpd + 1] - ffd[ltpd]; ? fd
```

```
Global`fd
```

```

fd[39] = 0.00074625
fd[40] = 0.0010712
fd[41] = 0.00115789
fd[42] = 0.00125725
fd[43] = 0.00137316
fd[44] = 0.00150651
fd[45] = 0.00166011
fd[46] = 0.00183575
fd[47] = 0.00203518
fd[48] = 0.00225907
fd[49] = 0.00250904
fd[50] = 0.00278564
fd[51] = 0.00308834
fd[52] = 0.00341655
fd[53] = 0.00376666
fd[54] = 0.00413699
fd[55] = 0.00452577
fd[56] = 0.00493502
fd[57] = 0.00536661

```

fd[58] = 0.00582317
fd[59] = 0.00630711
fd[60] = 0.0068215
fd[61] = 0.00737187
fd[62] = 0.00796328
fd[63] = 0.00860204
fd[64] = 0.00929735
fd[65] = 0.0100538
fd[66] = 0.0108802
fd[67] = 0.0117841
fd[68] = 0.0127737
fd[69] = 0.0138586
fd[70] = 0.0150505
fd[71] = 0.0163593
fd[72] = 0.0177975
fd[73] = 0.0193783
fd[74] = 0.021111
fd[75] = 0.0229948
fd[76] = 0.0250202
fd[77] = 0.0271677
fd[78] = 0.0294092
fd[79] = 0.0317032
fd[80] = 0.0339936
fd[81] = 0.0362083
fd[82] = 0.0382613
fd[83] = 0.0400502
fd[84] = 0.0414645
fd[85] = 0.0423843
fd[86] = 0.0426982
fd[87] = 0.0423112
fd[88] = 0.0411628
fd[89] = 0.0392382
fd[90] = 0.0365793
fd[91] = 0.0332876
fd[92] = 0.0295206
fd[93] = 0.0254764
fd[94] = 0.0213312
fd[95] = 0.017164
fd[96] = 0.0131716
fd[97] = 0.00961728
fd[98] = 0.00669733
fd[99] = 0.00446967
fd[100] = 0.0179512

La somme des fréquences des classes est égale à 1:

```
acc = 0; Do[acc = acc + fd[i], {i, zero, ltpd}]; acc
```

1.

■ Les quantiles

Des valeurs numériques de la fonction de distribution ont été calculées: **borne[i]** |--> **ffd[i]**

Les quantiles sont calculés sur la fonction réciproque: **ffd[i]** |--> **borne[i]**

Pour les valeurs intermédiaires, on effectue une interpolation linéaire.

A titre d'exemple, les quintiles sont calculés.

```
interpol[x0_, y0_, x1_, y1_, x_] := Module[{m = (y1 - y0) / (x1 - x0)}, y0 + m * (x - x0)]

Clear[qq]; qq[x_] := Module[{ind = zero}, While[ffd[ind] <= x, ind++];
  If[ind == zero, 0,
    interpol[ffd[ind - 1], borne[ind - 1], ffd[ind], borne[ind], x]]]

Table[qq[t], {t, 1/5, 4/5, 1/5}]

{72.4282, 80.3836, 85.5441, 90.4173}
```

■ L'histogramme

fd[i] = (probabilité de décès entre les âges **borne[i]** et **borne[i+1]**) est représenté par **l'aire** du rectangle.

h[i] = (probabilité **annuelle** de décès entre les âges **borne[i]** et **borne[i+1]**)

= (**densité** de décès entre les âges **borne[i]** et **borne[i+1]**) est représenté par **la hauteur** du rectangle.

```
Clear[h, i]; Do[h[i] = fd[i] / (borne[i + 1] - borne[i]), {i, zero, ltpd}]; ?h
```

```
Global`h
```

```
h[39] = 0.000995
```

```
h[40] = 0.0010712
```

```
h[41] = 0.00115789
```

```
h[42] = 0.00125725
```

```
h[43] = 0.00137316
```

```
h[44] = 0.00150651
```

```
h[45] = 0.00166011
```

```
h[46] = 0.00183575
```

```
h[47] = 0.00203518
```

```
h[48] = 0.00225907
```

```
h[49] = 0.00250904
```

```
h[50] = 0.00278564
```

```
h[51] = 0.00308834
```

```
h[52] = 0.00341655
```

```
h[53] = 0.00376666
```

```
h[54] = 0.00413699
```

```
h[55] = 0.00452577
```

h[56] = 0.00493502
h[57] = 0.00536661
h[58] = 0.00582317
h[59] = 0.00630711
h[60] = 0.0068215
h[61] = 0.00737187
h[62] = 0.00796328
h[63] = 0.00860204
h[64] = 0.00929735
h[65] = 0.0100538
h[66] = 0.0108802
h[67] = 0.0117841
h[68] = 0.0127737
h[69] = 0.0138586
h[70] = 0.0150505
h[71] = 0.0163593
h[72] = 0.0177975
h[73] = 0.0193783
h[74] = 0.021111
h[75] = 0.0229948
h[76] = 0.0250202
h[77] = 0.0271677
h[78] = 0.0294092
h[79] = 0.0317032
h[80] = 0.0339936
h[81] = 0.0362083
h[82] = 0.0382613
h[83] = 0.0400502
h[84] = 0.0414645
h[85] = 0.0423843
h[86] = 0.0426982
h[87] = 0.0423112
h[88] = 0.0411628
h[89] = 0.0392382
h[90] = 0.0365793
h[91] = 0.0332876
h[92] = 0.0295206
h[93] = 0.0254764
h[94] = 0.0213312
h[95] = 0.017164
h[96] = 0.0131716
h[97] = 0.00961728
h[98] = 0.00669733
h[99] = 0.00446967

Espérance de vie

Pour l'âge donné, l'âge moyen du décès est

$$\sum_{j=\text{zero}}^{\text{ltpd}} c[j] * \text{fd}[j]$$

où **c[j]** désigne le centre de la classe d'âge de **borne[j]** à **borne[j+1]**.

En soustrayant l'âge initial, on obtient l'espérance de vie

```
Clear[esp]; esp = -age +  $\sum_{j=\text{zero}}^{\text{ltpd}} \left( \frac{\text{borne}[j] + \text{borne}[j + 1]}{2} * \text{fd}[j] \right)$ 
```

41.7745