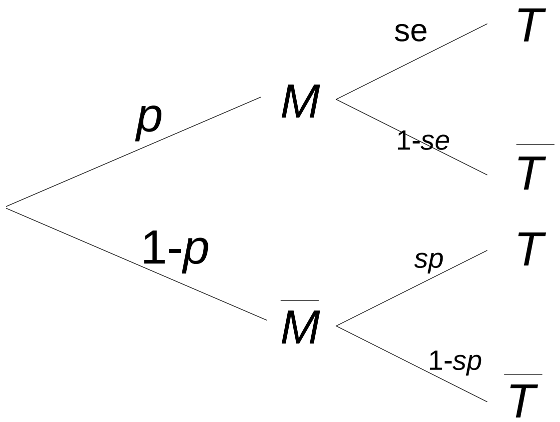
Les tests de dépistage – courbe ROC

Pour tout test servant à dépister une maladie, le fabriquant nous donne :

- la probabilité pour qu'un individu malade ait un test positif. Cette probabilité appelée sensibilité sera notée se.
- la probabilité pour qu'un individu sain ait un test négatif. Cette probabilité appelée spécificité sera notée sp.

Utilisons maintenant ce test sur une population dont on connaît la fréquence ou *prévalence* de la maladie qui sera noté ici *p*.

Le but de ce qui suit est de calculer la *valeur prédictive positive* (VPP) du test, c'est à dire la probabilité d'être malade, sachant le test est positif. C'est principalement ce résultat qui intéresse les médecins.



	T	$\overline{T}$	Total
M	$P(M\cap T)=p\times se$	$P(M\cap\overline{T})=p\times(1-se)$	P (M)
$\overline{M}$	$P(\overline{M}\cap T)=(1-p)\times sp$	$P(\overline{M}\cap\overline{T})=(1-p)(1-sp)$	$P(\overline{M})$
Total	P (T)	$P(\overline{T})$	1

La probabilité d'avoir un test positif est d'après la formule des probabilités totales :

$$P(T)=P(M\cap T)+P(\overline{M}\cap T)=p\times se+(1-p)(1-sp).$$

La probabilité d'être malade sachant que le test est positif est donc dans ces conditions :

$$VPP=P_T(M)=\frac{P(M\cap T)}{P(T)}=\frac{p\times se}{p\times se+(1-p)(1-sp)}.$$

Exemple : p = 0,02 ; $se = sp$ = 0,99. $P_T(M)=\dfrac{0{,}02\times 0{,}99}{0{,}02\times 0{,}99+0{,}98\times 0{,}01}\approx 0{,}669.$ Si la maladie est rare (prévalence faible), un test positif ne permet pas de conclure avec un bon niveau de confiance que le patient est bien malade. Néanmoins la probabilité d'avoir la maladie a été multipliée par 30.

De même la valeur prédictive négative sera :

$$VPN=P_{\overline{T}}(\overline{M})=\frac{P(\overline{M}\cap\overline{T})}{P(\overline{T})}=\frac{(1-p)(1-sp)}{p\times(1-se)+(1-p)(1-sp)}.$$

Dans l'exemple ci-dessus :

$$VPN = P_{\overline{T}}(\overline{M}) = \frac{0{,}98\times 0{,}01}{0{,}02\times 0{,}01+0{,}98\times 0{,}01} \approx 0{,}98.$$ On remarque que la VPN est presque égale à 1. C'est important car un médecin, à la vue d'un test négatif peut garantir avec un bon niveau de confiance que le patient n'est pas atteint de la maladie.

Remarque : la situation précédente peut être aussi représentée par un tableau

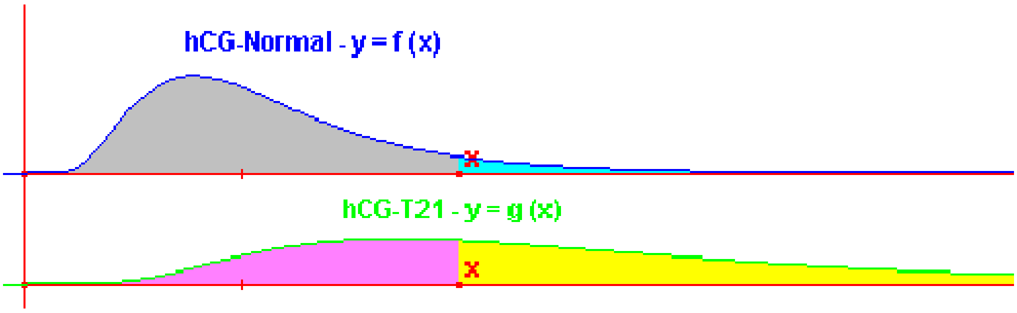
	T	$1-\overline{T}$	Total
M	$0{,}02\times 0{,}99=0{,}0198$	$0{,}02\times 0{,}01=0{,}0002$	0,02
$\overline{M}$	$0{,}98\times 0{,}01=0{,}0098$	$0{,}98\times 0{,}99=0{,}9702$	0,98
Total	0,0296	0,9704	1

La valeur diagnostique de ce test est dans ces conditions $\frac{0{,}0198}{0{,}0296}\approx 0{,}669$.

La page http://www.statistix.fr/IMG/t21/tm.html permet de voir l'influence de la sensibilité et de la spécificité sur la valeur diagnostique positive.

Application de cette situation à la mesure de la concentration de l'hormone HCG chez la femme enceinte.

f est la densité de la distribution de la concentration d'hCG chez la mère d'un enfant non trisomique. La densité est *g* si l'enfant est trisomique. On peut remarquer que la concentration d'hCG est supérieure chez la mère d'un trisomique. Utiliser le dosage d'hCG dans un contexte de test de dépistage revient donc à fixer un seuil *x*. Le test sera considéré comme positif si la concentration d'hCG dépasse *x*.



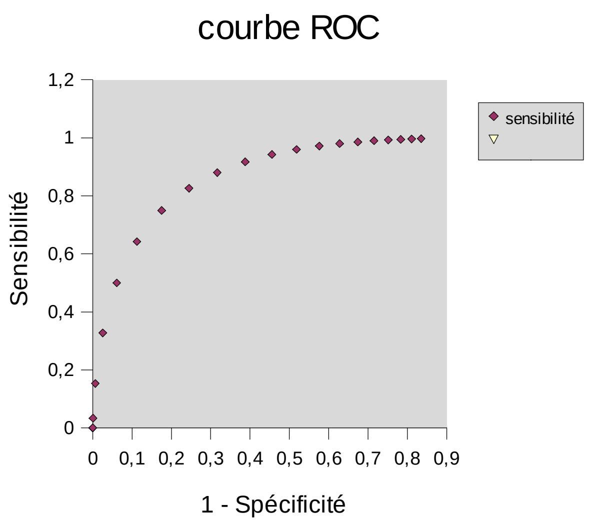
L'aire de couleur grise sous la courbe de la fonction *f* est la probabilité pour que le test soit négatif sachant que l'enfant n'est pas trisomique, c'est donc la *spécificité*. On a intérêt à la prendre la plus grande possible. De même la probabilité que le test soit positif sachant que l'enfant est trisomique est représentée par l'aire de couleur jaune, c'est la *sensibilité*. On peut remarquer que plus la spécificité augmente plus la sensibilité diminue. Dans le modèle utilisé à la page http://www.statistix.fr/IMG/t21/hcg.htm une spécificité de 0,95 revient à choisir un seuil de 2,275 ce qui implique une sensibilité égale à 0,5203.

x	Spécificité	1-spé.	Sensibilité
0,0	1,0000	0,0000	0,0000
0,2	0,9999	0,0001	0,0007
0,4	0,9993	0,0007	0,0335
0,6	0,9933	0,0067	0,1535
0,8	0,9744	0,0256	0,3277
1,0	0,9388	0,0612	0,5000
1,2	0,8876	0,1124	0,6424
1,4	0,8247	0,1753	0,7496
1,6	0,7551	0,2449	0,8264
1,8	0,6832	0,3168	0,8802
2,0	0,6122	0,3878	0,9172
2,2	0,5445	0,4555	0,9426
2,4	0,4815	0,5185	0,9601
2,6	0,4239	0,5761	0,9720
2,8	0,3719	0,6281	0,9803
3,0	0,3256	0,6744	0,9860
3,2	0,2845	0,7155	0,9900
3,4	0,2483	0,7517	0,9929
3,6	0,2166	0,7834	0,9948
3,8	0,1889	0,8111	0,9963
4,0	0,1647	0,8353	0,9973

courbe ROC (Receiver Operating Characteristic)

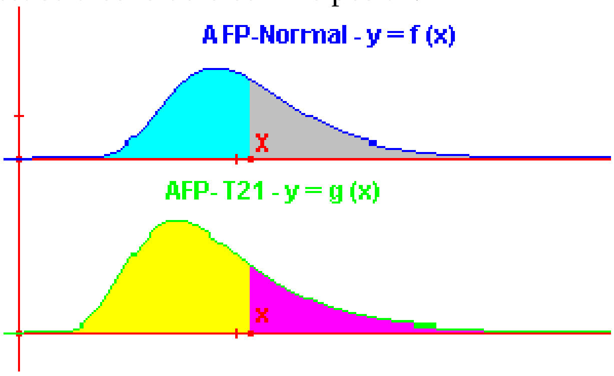
Le tableau suivant donne pour plusieurs valeurs de x la sensibilité et spécificité du test à partir du modèle utilisé à la page http://www.statistix.fr/IMG/t21/hcg.htm.

Faire une « courbe ROC » consiste à faire varier x et à représenter dans un plan les points de coordonnées $se = P_M(T)$ et $1 - p_{sp} = P_{\overline{M}}(T)$.



Comparaison avec la mesure de la concentration de l'hormone AFP

Comme précédemment, f est la densité de la distribution de la concentration d'AFP chez la mère d'un non trisomique. La densité est g si l'enfant est trisomique. On peut remarquer que la concentration d'AFP est inférieure chez la mère d'un trisomique et donc à l'inverse de la situation précédente, utiliser le dosage d'AFP dans un contexte de test de dépistage revient donc à fixer un seuil x en deçà duquel le test sera considéré comme positif.



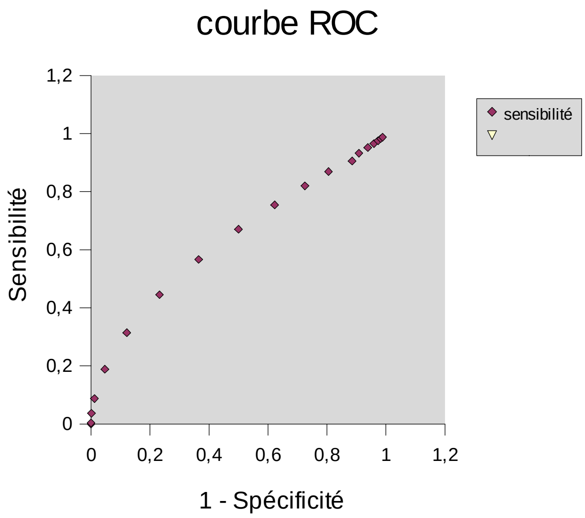
x étant donné, l'aire de couleur grise sous la courbe de la fonction f sera donc la probabilité pour que le test soit négatif sachant que l'enfant est trisomique (*spécificité*). De même la *sensibilité* est représentée par l'aire de couleur jaune, sous la courbe de g. Comme pour le test précédent, on peut remarquer que plus la spécificité augmente, plus la sensibilité diminue. Dans le modèle utilisé à la page http://www.statistix.fr/IMG/t21/afp.htm une spécificité de 0,95 revient à choisir un seuil de 0,6055. La sensibilité est dans ce cas 0,1949. Elle est beaucoup plus faible que pour l'hCG.

x	Spécificité	1-spé.	Sensibilité
0,0	1,0000	0,0000	0,0000
0,1	0,9999	0,0001	0,0001
0,2	0,9999	0,0001	0,0001
0,3	0,9999	0,0001	0,0037
0,4	0,9986	0,0014	0,0365
0,5	0,9884	0,0116	0,0875
0,6	0,9530	0,0470	0,1884
0,7	0,8788	0,1212	0,3142
0,8	0,7677	0,2323	0,4452
0,9	0,6351	0,3649	0,5666
1,0	0,5000	0,5000	0,6704
1,1	0,3773	0,6227	0,7544
1,2	0,2749	0,7251	0,8196
1,3	0,1948	0,8052	0,8690
1,4	0,1349	0,8651	0,9056
1,5	0,0918	0,9082	0,9323
1,6	0,0616	0,9384	0,9516
1,7	0,0409	0,9591	0,9655
1,8	0,0269	0,9731	0,9754
1,9	0,0176	0,9824	0,9824
2,0	0,0115	0,9885	0,9875

courbe ROC

Le tableau ci-contre donne pour plusieurs valeurs de *x* la sensibilité et spécificité du test à partir du modèle utilisé à la page http://www.statistix.fr/IMG/t21/afp.htm.

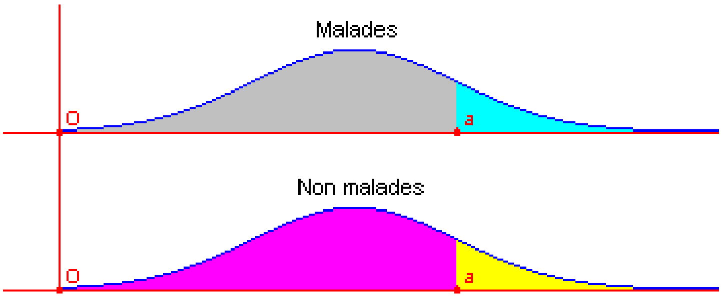
D'où la courbe ROC :



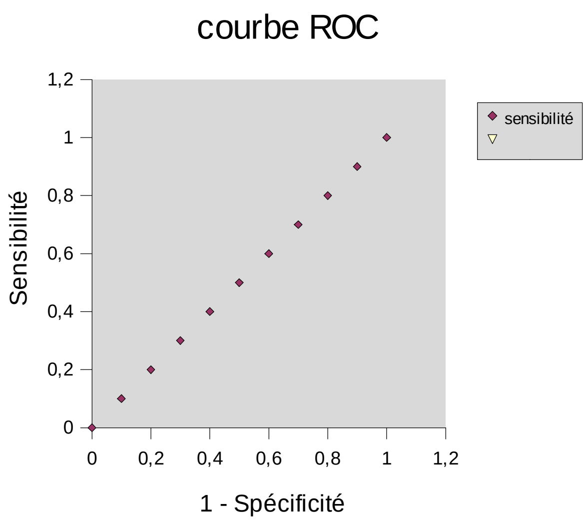
la courbe est beaucoup plus « plate » que précédemment, cela signifie que le test est moins bon.

Les ROC dans deux cas extrêmes fictifs

Imaginons que la concentration d'une certaine hormone ait la même distribution chez les malades et les non malades :

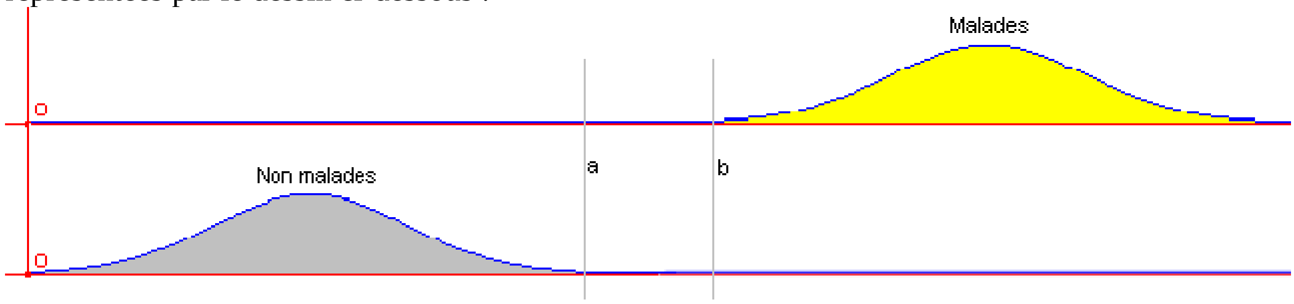


Il est clair que, dans cette situation, la concentration de l'hormone est indépendante de la maladie mais imaginons quand même que l'on met en place un test basé sur la concentration de cette hormone. On fixera un seuil et on déclarera que le test est positif si on est au delà de ce seuil, mais comme on ne peut distinguer les distributions on peut faire aussi le contraire : déclarer que le test est positif si on est en deçà de ce seuil !

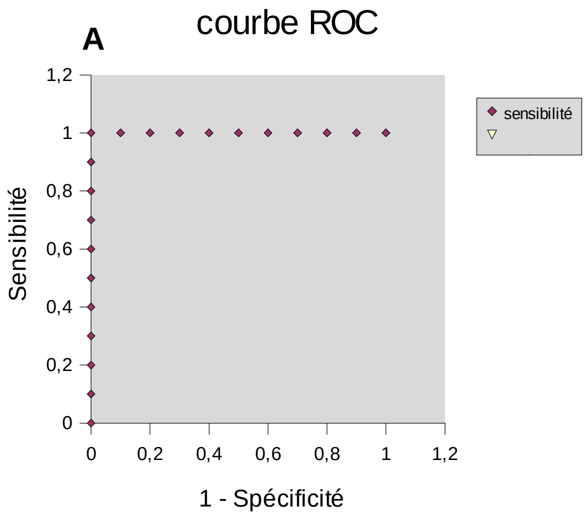


De toute façon, dans les deux cas, la spécificité et la sensibilité du test ont pour somme 1 car les distributions de la concentration étant identiques chez les malades et les non-malades, les aires représentant la spécificité et la sensibilité sont complémentaires. D'où l'alignement des points de la courbe ROC :

A l'inverse, imaginons maintenant que les densités de la concentration de l'hormone soient représentées par le dessin ci-dessous :



Le test sera déclaré positif si la concentration de l'hormone dépasse un seuil *x*. Si *x* est inférieur à *a*, la spécificité représentée par l'aire de couleur grise variera de 0, si *x* = 0, à 1 si *x* = *a* alors que la sensibilité représentée par l'aire jaune vaudra 1. Si *x* est compris entre *a* et *b* alors la spécificité et la sensibilité valent 1, et si *x* dépasse *b* la spécificité reste égale à 1 alors que la sensibilité diminuera de 1 à 0. Dans cette situation on aura bien-sûr intérêt à prendre *x* compris entre *a* et *b*.



Ci-contre la courbe ROC obtenue dans cette situation. *x* compris entre *a* et *b* correspond au point A de la courbe ROC.

En conclusion, plus la courbe ROC se rapprochera de la courbe ci-contre meilleur sera le test. Dans le cas contraire, si les points sont presque alignés (dessin ci-dessus) le test sera moins puissant, c'est à dire peu apte à séparer le groupe de malades du groupe des non-malades.