2 Probabilités conditionnelles. Événements indépendants

2.1 Probabilité conditionnelle

Soient A et B deux événements tels que $P(B) > 0$. Soit alors $P(A \mid B)$, la probabilité que A se réalise, B étant réalisé. Dans la mesure où nous savons que B s'est réalisé, B devient le nouvel espace d'échantillonnage, remplaçant ainsi Ω, ce qui conduit à la définition suivante :

$$P(A \mid B) = \frac{P(AB)}{P(B)}$$

$P(A \mid B)$ est appelée *probabilité conditionnelle* de A, B étant donné, c'est-à-dire la probabilité de réalisation de A, étant admis que B s'est réalisé. Il est facile de montrer, que la probabilité conditionnelle obéit aux axiomes énoncés. De plus, on a en particulier :

- $P(A \mid B) = 1$ si $A \supset B$
- $P(A \mid B) = 0$ si $A \cdot B = \emptyset$.

Relativement à l'événement A, on dit parfois que l'événement B fournit une information supplémentaire positive (ou négative) si on a $P(A \mid B) > P(A)$ (ou $P(A \mid B) < P(A)$).

Exemple: *On lance un dé équilibré ; soient les événements* $A = \{\textit{nombre pair}\}$ *est* $B = \{\textit{nombre} \geq 4\}$. $P(A \mid B) = ?$

Réponse: Alors $P(A \mid B) = 2/3$ tandis que $P(A) = \frac{1}{2}$. Le fait de savoir que l'événement B est réalisé constitue donc ici une information supplémentaire positive en ce qui concerne l'événement A dont elle accroît la probabilité.

Problème: Un examen comprend une partie théorique (réussite $= A$) et une partie pratique (réussite $= B$). Au cours d'une année, 500 personnes se sont présentées à cet examen, dont

- 200 ont réussi les deux parties,
- 100 ont réussi seulement la partie théorique et
- 50 ont réussi seulement la partie pratique.

Quelle est la probabilité qu'une personne choisie au hasard ait réussi la partie pratique ? De même, si on sait qu'elle a déjà passé avec succès la partie théorique ?

On verra que la probabilité conditionnelle constitue un excellent moyen d'évaluer le degré de dépendance stochastique existant entre deux événements du même ensemble fondamental Ω.

2.2 Théorème de multiplication

Quels que soient deux événements A, B utilisant la probabilité conditionnelle on peux obtenir que,

$$P(AB) = P(A \mid B) \cdot P(B) = P(B \mid A) \cdot P(A)$$

pourvu que $P(A) > 0$, $P(B) > 0$.

Cette réalisation se géneralise facilement à plus de deux événements :

$$P(A_1 A_2 \ldots A_n) = P(A_1) P(A_2 \mid A_1) \cdot P(A_3 \mid A_1 \cdot A_2) \ldots P(A_n \mid A_1 \cdot A_2 \ldots A_n)$$

si tous les événement $A_1, A_2, \ldots, A_n$ sont de probabilité non nulle.

Exemple: (contrôle de qualité) *D'un lot comprenant 10 pièces et dont la moitié sont défectueuses, on prélève (sans remise des pièces tirées) un échantillon de taille* 3. *Quelle est la probabilité $P(A)$ que l'échantillon ne comprenne aucune pièce défectueuse ?*

Réponse: Soit $A_k = \{$ l'événement que la $k^{\text{ième}}$ pièce prélevée soit bonne $\}$. Alors

$$P(A) = P(A_1 \cdot A_2 \cdot A_3) = \frac{5}{10} \cdot \frac{4}{9} \cdot \frac{3}{8} = \frac{1}{12}.$$

2.3 Théorème des probabilités totales

Soient A et B deux événements quelconques. Nous pouvons écrire A sous la forme $A = AB + A\overline{B}$, car tout élément de A doit se trouver soit dans A et B à la fois, soit dans A mais pas dans B.

Comme évidemment AB et $A\overline{B}$ s'excluent mutuellement, on peut utiliser l'axiome :

$$P(A) = P(AB) + P(A\overline{B}) = P(A \mid B) P(B) + P(A \mid \overline{B}) P(\overline{B}).$$

Exemple: *Dans l'exemple ci-dessus, quelle est la probabilité que la deuxième pièce prélevée soit bonne ?*

Réponse:

$$P(A_2) = P(A_2 \mid A_1) P(A_1) + P(A_2 \mid \overline{A}_1) P(\overline{A}_1) = \frac{4}{9} \cdot \frac{5}{10} + \frac{5}{9} \cdot \frac{5}{10} = \frac{1}{2}.$$

L'équation peut être interprétée de la façon suivante : la probabilité de l'événement A est une moyenne pondérée de la probabilité conditionnelle de A lorsque B est apparu et de la probabilité conditionelle du même A lorsque B n'est pas apparu, les poids étant les probabilités des événements conditionnants. Il existe de nombreuses situations où il est difficile de calculer directement la probabilité d'un événement mais où il est par contre possible de la calculer connaissant ses probabilités conditionnelles si certains événements sont réalisés.

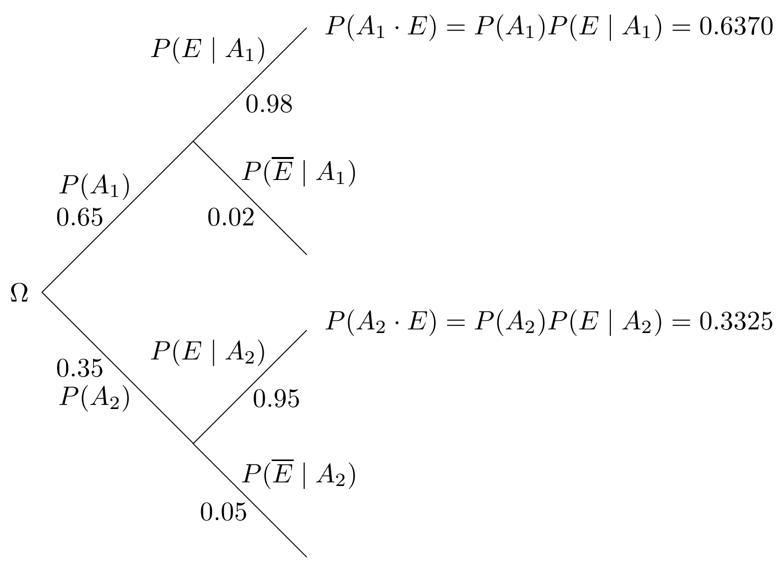
Exemple: *Dans une usine deux machines A_1, A_2 fabriquent des boulons de même type. A_1 produit en moyenne 2% de boulons défectueux $(\overline{E})$, A_2 5%. A_1 fabrique 65% de la production totale. On choisit un boulon au hasard. Quelle la probabilité qu'il n'est pas défectueux (E).*

Réponse: La probabilité demandée est

$$P(E) = P(E \mid A_1)P(A_1) + P(E \mid A_2)P(A_2) = 0.9695,$$

selon la formule des probabilités totales.

Remarque: On peut ainsi illustrer cela :



Exemple: *Une compagnie d'assurance estime que les gens peuvent être répartis en deux classes : ceux qui sont enclins aux accidents et ceux qui ne le sont pas. Ses statistiques montrent qu'un individu enclin aux accidents a une probabilité de 0.4 d'en avoir un dans l'espace d'un an; cette probabilité tombe à 0.2 pour les gens à* *risque modéré. On suppose que 30% de la population appartient à la classe à haut risque. Quelle est alors la probabilité qu'un nouvel assuré soit victime d'un accident durant l'année qui suit la signature de son contrat ?*

Réponse: Nous obtiendrons la probabilité de l'événement cité en le conditionnant selon que la signature de la police est ou n'est pas enclin aux accidents. On note par

$$A_1 = \{\text{le signataire aura un accident dans l'année qui suit l'établissement du contrat}\}$$

et par

$$A = \{\text{ le signataire est enclin aux accidents}\}.$$

La probabilité voulue est alors donnée par

$$P(A_1) = P(A_1 \mid A)P(A) + P(A_1 \mid \overline{A}) \cdot P(\overline{A}) = 0.4 \cdot 0.3 + 0.2 \cdot 0.7 = 0.26.$$

Considérons un *système exhaustif d'événements*, c'est-à-dire un ensemble au maximum infinie dénombrable $\{A_1, A_2, \ldots\}$ d'événements deux à deux incompatibles, tel que $\sum A_i = \Omega$. La probabilité d'un événement quelconque B peut alors être calculée à partir des probabilités conditionnelles $P(B \mid A_i)$. On a $P(B) = \sum P(A_i)P(B \mid A_i)$. $P(B)$ est la probabilité de B indépendamment de la réalisation d'un quelconque A_i. Pour souligner la différence de la notion de probabilité conditionnelle, on parle de *probabilité totale* de B.

2.4 Théorème de Bayes

La formule de Bayes est simplement derivée de probabilité totale.

Si $\{A_1, A_2, \ldots, A_n\}$ est un système exhaustif d'événements, et si B est un événement quelconque,

$$P(A_k \mid B) = \frac{P(B \mid A_k) \cdot P(A_k)}{\sum_{i=1}^{n} P(B \mid A_i) \cdot P(A_i)} \qquad \text{pour } k = 1, \ldots, n.$$

L'équation est appelée *formule de Bayes*, du nom du philosophe anglais Thomas Bayes. Si nous traitons les événements A_k comme les hypothèses possibles sur une question donnée, la règle du Bayes joue un rôle utile en nous montrant comment les opinions à priori sur ces hypothèses doivent être modifiées à la lumière du résultat de l'expérience.

Ce théorème permet d'évaluer les probabilités des différents événements $A_1, A_2, \ldots, A_n$ qui peuvent causer la réalisation de B. Pour cette raison le théorème ou la règle de Bayes est souvent désignée comme un théorème sur la *probabilité des causes.*

Exemple: *Une entreprise utilise trois types d'ampoules électriques notés T_1, T_2, T_3, dans les proportions 60%, 30%, 10%. Les probabilités de bon fonctionnement de ces trois types pour un temps donné s'élèvent à 0.9, 0.8 et 0.5 respectivement. Quelle est la probabilité qu'une ampoule tombée en panne soit du type T_1 ?*

Réponse: Si on introduit les événements

$$\begin{aligned} B &= \{ \text{ une ampoule choisie au hasard tombe en panne } \} \\ A_k &= \{ \text{ une ampoule est du type } T_k\} \text{ on a} \end{aligned}$$

$$P(A_1 \mid B) = \frac{P(B \mid A_1) \cdot P(A_1)}{\sum_{i=1}^{3} P(B \mid A_i) \cdot P(A_i)} = \frac{0.1 \cdot 0.6}{0.1 \cdot 0.6 + 0.2 \cdot 0.3 + 0.5 \cdot 0.1} = \frac{6}{17}.$$

Exemple: *Pour dépister une maladie, on applique un test. Si le patient est effectivement atteint, le test donne un résultat positif dans 99% des cas. Mais il se peut aussi que le résultat du test soit positif alors que le consultant est en bonne santé, et ceci se produit dans 2% des cas. Sachant qu'en moyenne un consultant sur 1000 est atteint de la maladie à dépister, calculer la probabilité pour qu'un client soit atteint sachant que son test a été positif. Comment améliorer ce résultat ?*

Réponse: Soit M l'événement "le client est atteint" et T_+ l'événement "le test est positif". On cherche à calculer $P(M \mid T_+)$. Les données sont

$$P(M) = 0.001, \quad P(T_+ \mid M) = 0.99, \quad P(T_+ \mid \overline{M}) = 0.02.$$

D'après la définition de Bayes,

$$P(M, T_+) = P(M \mid T_+)P(T_+) = P(T_+ \mid M)P(M),$$

d'où

$$P(M \mid T_+) = \frac{P(T_+ \mid M)P(M)}{P(T_+)}.$$

La règle des causes totales nous donne d'autre part

$$P(T_+) = P(T_+ \mid M)P(M) + P(T_+ \mid \overline{M})P(\overline{M}),$$

d'où

$$P(M \mid T_+) = \frac{P(T_+ \mid M)P(M)}{P(T_+ \mid M)P(M) + P(T_+ \mid \overline{M})P(\overline{M})}.$$

Le résultat numérique est

$$P(M \mid T_+) \sim \frac{(0.99)(0.001)}{(0.99)(0.001) + (0.02)(0.999)} = \frac{1}{21},$$

ce qui est très faible ! On voit qu'il faut un $P(T_+ \mid \overline{M})$ plus faible.

Par exemple si $P(T_+ \mid \overline{M}) = 0.002$ on trouve $P(M \mid T_+) = \frac{1}{3}$, et si $P(T_+ \mid \overline{M}) = 0.0002$ on trouve $P(M \mid T_+) = \frac{99}{101}$.

Commentaire. Dans la pratique des dispensaires, plutôt que de faire un test avec faible $P(T_+ \mid \overline{M})$ qui risque de coûter cher, on refait passer un test plus sûr aux patients ayant eu un premier test positif. L'essentiel est que le premier test ne laisse pas passer un malade atteint. On doit donc avoir un $P(T_+ \mid M)$ très proche de 1.

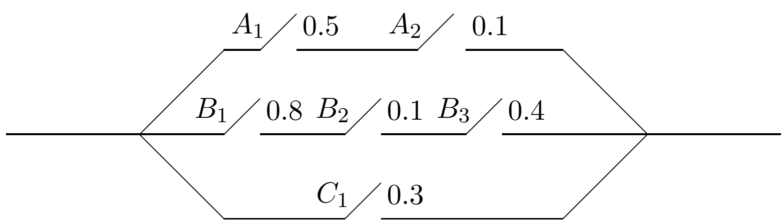
La situation décrite plus haut est celle de certains tests de dépistage de la maladie. Le premier test en usage dans les dispensaires a un $P(T_+ \mid \overline{M})$ relativement élevé : entre autres raisons, une certaine forme de grippe donne un test positif.

2.4.1 Événements indépendants

La notion de l'indépendance joue un rôle de premier ordre en calcul des probabilités. Intuitivement, on dira que A et B sont indépendants si la probabilité d'apparition d'un événement n'est pas influencée par la réalisation (ou non-réalisation) de l'autre.

Le concept de probabilité conditionnelle permet de préciser cette idée intuitive. Si $P(A \mid B) = P(A)$, c'est-à-dire si la probabilité de réalisation de A n'est pas affectée par la réalisation (ou le non-réalisation) de B, nous disons que A et B sont des événements indépendants. Ce résultat est équivalent à : $P(A \cdot B) = P(A) \cdot P(B)$. Inversement, si $P(A) \cdot P(B) = P(AB)$ est satisfaite, A est B sont indépendants.

Exemple: *Considérez le circuit suivant :*



Les A, B, C sont des relais, en position ouverte ou fermée. Les nombres indiquent la probabilité que le relais correspondant soit ouvert. Les relais sont "indépendants" (formalisez cette notion).

Calculez la probabilité pour que le cricuit total "passe" (c'est-à-dire qu'il existe au moins une branche sur laquelle tous les relais sont fermés).

Réponse: Appelons X_i l'événement "A_i est ouvert", Y_i l'événement "B_i est ouvert" et Z_i l'événement "C_i est ouvert". Les données sont :

$$\begin{cases} P(X_1) = 0.5 & P(X_2) = 0.1 \\ P(Y_1) = 0.8 & P(Y_2) = 0.1 \quad P(Y_3) = 0.4 \\ P(Z_1) = 0.3. \end{cases}$$

A cela il faut ajouter la donnée qualitative : les relais sont indépendants. Dire que les relais sont indépendants c'est dire qu'on a toutes sortes d'égalités du type

$$P(\overline{X}_1 \cap \overline{X}_2) = P(\overline{X}_1)P(\overline{X}_2),$$

et de même si on appelle P_i l'événement "la branche i passe", on aura par exemple $P(\overline{P}_1 \cap \overline{P}_2 \cap \overline{P}_3) = P(\overline{P}_1)P(\overline{P}_2)P(\overline{P}_3)$.

Comme $P_1 = \overline{X}_1 \cap \overline{X}_2$, on a $P(P_1) = P(\overline{X}_1)P(\overline{X}_2)$. De même $P(P_2) = P(\overline{Y}_1)P(\overline{Y}_2)P(\overline{Y}_3)$ et $P(P_3) = P(\overline{Z}_1)$.

Si K est l'événement "le circuit passe" alors $\overline{K} = \overline{P}_1 \cap \overline{P}_2 \cap \overline{P}_3$ si bien que $P(K) = 1 - P(\overline{K}) = 1 - P(\overline{P}_1)P(\overline{P}_2)P(\overline{P}_3)$ et donc

$$P(K) = 1 - (1 - P(\overline{X}_1)P(\overline{X}_2))(1 - P(\overline{Y}_1)P(\overline{Y}_2)P(\overline{Y}_3))(1 - P(\overline{Z}_1)).$$

Numériquement : $P(K) = 0.95094$.

2.4.2 Indépendance entre plusieurs événements

Supposons maintenant que A soit indépendant de B et de C. Est-ce que A le sera de BC ? Aussi surprenant que cela puisse paraître, la réponse est non, comme le montre l'exemple suivant.

Exemple: *On jette deux dés équilibrés. Soient les événements*

$A = \{$ *la somme est* $7\}$,
$B = \{$ *le preimer dé montre* $4\}$
$C = \{$ *le second dé donne* $3\}$.

On peux montrer que A et B sont indépendants et que A et C le sont aussi. Cependant A n'est manifestement pas indépendant de BC puisque $P(A \mid BC) = 1$. On comprend grâce à cet exemple qu'une bonne définition de l'indépendance de trois événements ne peut pas se limiter à exiger que les événements soient indépendants par paires dans les $\binom{3}{2}$ combinaisons possibles. Trois événements A, B, C sont dits totalement indépendants si

- $P(ABC) = P(A)P(B)P(C)$
- $P(AB) = P(A)P(B)$
- $P(AC) = P(A)P(C)$
- $P(BC) = P(B)P(C)$.

Il est évidemment possible d'étendre la définition d'indépendance totale à plus de trois événements : un ensemble d'événements $A_,A_2, \ldots, A_n$ est dit totalement indépendant si pour tout sous-ensemble $A_{1'}, A_{2'}, \ldots, A_{k'}$, $k \leq n$.

$$P(A_{1'} \cdot A_{2'} \cdot \ldots A_{k'}) = P(A_{1'})P(A_{2'}) \ldots P(A_{k'}).$$

Généralement on a besoin de $2^n - n - 1$ relations pour définir indépendance de n événements.

2.4.3 Remarques

La définition de l'indépendance de deux événements suscite plusieurs remarques.

- On vérifie facilement que l'indépendance de A et B implique celle de A et $\overline{B}$; de $\overline{A}$ et B ; de $\overline{A}$ et $\overline{B}$.
- Deux événements incompatibles et de probabilité non nulle sont toujours dépendants.
- Si $B \subset A$ alors les événement A, B ne sont pas indépendants.

2.5 Modèles d'urne

2.5.1 Différents modes de tirage

Dans de nombreuses situations, une expérience stochastique s'effectue en plusieurs étapes qui peuvent être regardées comme des expériences partielles. Pour l'étude du cas particulier où ces expériences partielles sont indépendantes les unes des autres, il est souvent avantageux de faire appel aux *modèles d'urne*. Considérons donc une urne contenant N boules distinctes. De cette urne on extrait au hasard n boules. Cette opération peut se faire de plusieurs manières différentes :

– On peut tirer les boules l'une après l'autre, en remettant chaque fois la boule tirée avant de procéder au tirage suivant. On parle alors d'un *tirage avec remise* ou *non exhaustif.*
– On peut tirer les boules l'une après l'autre, sans qu'une boule tirée soit remise dans l'urne. On parle alors d'un *tirage sans remise* ou *exhaustif.*
– On peut tirer les n boules *simultanément*, c'est-à-dire d'un seul coup.

2.5.2 Urne contenant deux sortes de boules

Supposons maintenant que le contenu d'une urne soit composé de r boules blanches et de $N-r$ boules noires. Nous nous intéressons à la répartition des deux couleurs dans un échantillon de taille n prélevé de cette urne, c'est-à-dire à la probabilité $P(A_k)$ où $A_k = \{$ l'échantillon comprend exactement k boules blanches $\}$. Nous considérons séparément chacune des trois méthodes de tirage présentées ci-dessus.

– Lorsque les boules sont prélevées *avec remise*, on a affaire à une succession de n expériences partielles qui sont indentiques et indépendantes l'une de l'autre. On obtient alors :

$$P(A_k) = \binom{n}{k} p^k (1-p)^{n-k} \text{ où } p = r/N \text{ est la probabilité}$$

qu'une boule tirée soit blanche ($k = 0, 1, \ldots, n$).
– Lorsqu'on effectue un tirage de n boules *sans remise*, les n expériences partielles dont est composé ce processus, à savoir les prélèvement des n boules, ne sont ni identiques, ni indépendantes l'une de l'autre. En effet la probabitité de prélever une boule blanche varie constamment au cours du tirage et dépend des résultats déjà obtenus. En faisant appel au théorème de multiplication, on peut démontrer que la probabilité d'avoir exactement k boules blanches parmi les n boules tirées est donnée par

$$P(A_k) = \frac{\binom{r}{k}\binom{N-r}{n-k}}{\binom{N}{n}} \qquad \text{où } k = 0, 1 \ldots \min(r, n).$$

– Lorsqu'on tire les n boules *simultanément*, on peut modéliser ce phénomène stochastique par une loi uniforme discrète. Les calculs du nombre de cas possibles et du nombre de cas favorables à la réalisation de A_k montrent que la probabilité de cet événement est la même que celle obtenue pour le tirage sans remise.

Remarque: On peut montrer que

$$\frac{\binom{r}{k}\binom{N-r}{n-k}}{\binom{N}{n}} \to \binom{n}{k} p^k (1-p)^{n-k}$$

si $N \to \infty$, $r \to \infty$ tel que $r/N \to p$. Ceci signifie que le modèle sans remise se rapproche du modèle avec remise si l'effectif N de l'urne est grand par rapport à la taille n de l'échantillon.

Exemple: *Un article est produit en masse ; on sait qu'en moyenne 10% des pièces sont défectueuses. L'article est vendu dans des emballages contenant chacun dix pièces. Le fournisseur garantit qu'il y ait au moins huit pièces non-défectueuses dans chaque emballage. Quelle est la probabilité que cette garantie puisse être tenue ?*

Réponse: On peut assimiler cette situation à un tirage sans remise de dix boules d'une urne comprenant un nombre infiniment grand de boules dont 90% sont blanches. La probabilité que la garantie du fournisseur puisse être tenue se calcule alors par

$$P(A) = P(A_8 + A_9 + A_{10}) = \sum_{k=8}^{10} P(A_k) = \sum_{k=8}^{10} \binom{10}{k} 0.9^k \cdot 0.1^{10-k} = 0.9298,$$

où $A_k = \{\text{un emballage comprend } k \text{ bonnes pièces}\}$.

2.6 Application à des problèmes de fiabilité

La *théorie de la fiabilité* a comme objectif d'étudier l'aptitude de *dispositifs techniques* à accomplir une fonction requise, dans des conditions données et pendant une durée donnée.

Nous supposons que chaque dispositif se trouve dans l'un des deux états suivants :
– apte à fonctionner,
– inapte à fonctionner, c'est-à-dire hors service.

Nous définissons donc la *fiabilité* d'un dispositif comme étant la probabilité qu'il fonctionne correctement pendant un intervalle de temps donné, c'est-à-dire qu'il n'a pas de défaillance pendant cet intervalle.

2.6.1 Fiabilité des systèmes

La plupart des produits industriels représentent des *systèmes* formés d'un grand nombre de composants en interaction qui peuvent être regroupés en sous-systèmes. Pour évaluer la fiabilité d'un système, il faut donc connaître

– la fiabilité des éléments constituants
– la *structure de fiabilité* du système.

Nous considérons par la suite un certain nombre de systèmes ou de configurations simples. Si nous définissons les événements

$$\begin{aligned} S &= \{ \text{ le système fonctionne } \} \\ A_k &= \{ \text{ le } k^{\text{ième}} \text{ élément fonctionne } \} \text{ où } k = 1, 2, \ldots, n, \end{aligned}$$

le problème consiste à déterminer $P(S)$ en fonction des $P(A_k)$.

Un *système en série* ne fonctionne que si tous ses éléments fonctionnent. La défaillance d'un élément quelconque entraîne donc la défaillance du système. Il en résulte que $S = A_1 \cdot A_2 \cdot \ldots \cdot A_n$ d'où

$$P(S) = \prod_{k=1}^{n} P(A_k).$$

Un *système en parallèle* ne fonctionne que si au moins un de ses éléments fonctionne. Une défaillance du système ne se produit donc que si tous les éléments sont défaillants. Il en résulte que

$$S = A_1 + A_2 + \ldots + A_n \text{ d'où } P(S) = 1 - \prod_{k=1}^{n}(1 - P(A_k)).$$

Un *système k-de-n* ne fonctionne que si au moins k de ses éléments fonctionnent. Si nous admettons que tous les éléments ont la même fiabilité $P(A_k) = p$, il en résulte d'après

$$P(S) = \sum_{i=k}^{n} \binom{n}{i} p^i (1-p)^{n-i}.$$

Notons que les systèmes en série et les systèmes en parallèle peuvent être considérés comme des cas particuliers d'un système k-de-n.