Université Denis Diderot Paris 7

Compléments du cours de Probabilités

1 Théorème IV.1 du cours caractérisant les mesures de probabilité sur $\mathbb{R}$: énoncé et preuve

Rappelons l'énoncé du théorème :

Theorem 1.1. *Soit* $F : \mathbb{R} \to [0,1]$ *croissante, continue à droite, telle que*

$$\lim_{x\to-\infty} F(x) = 0, \quad \lim_{x\to\infty} F(x) = 1.$$

Alors il existe une unique mesure de proba $\mathbb{P}$ *définie sur* $\mathcal{B}(\mathbb{R})$ *telle que*

$$\forall x \in \mathbb{R} \quad \mathbb{P}((-\infty, x]) = F(x). \tag{1}$$

Il y a deux choses à démontrer : l'existence d'une telle mesure, et son unicité.

1.1 Unicité

Si on note

$$\mathcal{C} := \{(-\infty, x], x \in \mathbb{R}\}$$

les valeurs de $\mathbb{P}(C), C \in \mathcal{C}$ nous sont imposées par (1).
Il est facile de voir que cette condition implique que $\mathbb{P}$ est en fait imposée sur la classe plus large des intervalles de $\mathbb{R}$. En effet d'après les propriétés de mesure, on doit vérifier

$$\mathbb{P}((-\infty, b)) = \lim_{n\to\infty} \mathbb{P}((-\infty, b-1/n]) = \lim_{n\to\infty} F(b-1/n) =: F(b-).$$

Notons que la limite $F(b-)$ ci-dessus existe bel et bien, quelque soit $b \in \mathbb{R}$, puisque F est supposée croissante.
Notons par ailleurs que l'hypothèse de continuité à droite de F est nécessaire, puisqu'il faut que

$$F(b) = \mathbb{P}((-\infty, b]) = \lim_{n\to\infty} \mathbb{P}((-\infty, b+1/n]) = \lim_{n\to\infty} F(b+1/n).$$

On a aussi, en utilisant que $\lim_{x\to-\infty} F(x) = 0$, $\lim_{x\to\infty} F(x) = 1$,

$$\mathbb{P}(\emptyset) = \lim_{n\to\infty} \mathbb{P}((-\infty, -n]) = 0,$$

$$\mathbb{P}(\mathbb{R}) = \lim_{n\to\infty} \mathbb{P}((-\infty, n]) = 1,$$

$$\mathbb{P}((a, \infty)) = \mathbb{P}(\mathbb{R}) - \mathbb{P}((-\infty, a]) = 1 - F(a),$$

$$\mathbb{P}([a,\infty)) = \mathbb{P}(\mathbb{R}) - \mathbb{P}((-\infty,a)) = 1 - F(a-),$$

On a donc également, pour $a < b$

$$\mathbb{P}((a,b]) = \mathbb{P}((-\infty,b]) - \mathbb{P}((-\infty,a]) = F(b) - F(a),$$

$$\mathbb{P}((a,b)) = \mathbb{P}((-\infty,b)) - \mathbb{P}((-\infty,a]) = F(b-) - F(a),$$
$$\mathbb{P}([a,b]) = \mathbb{P}((-\infty,b]) - \mathbb{P}((-\infty,a)) = F(b) - F(a-),$$
$$\mathbb{P}([a,b)) = \mathbb{P}((-\infty,b)) - \mathbb{P}((-\infty,a)) = F(b-) - F(a-).$$
$$\mathbb{P}(\{a\}) = \mathbb{P}((-\infty,a]) - \mathbb{P}((-\infty,a)) = F(a) - F(a-).$$

Ainsi $\mathbb{P}$ est définie sur tout intervalle de $\mathbb{R}$. On notera que pour que les valeurs ci-dessus soient toujours positives, il est nécessaire que F soit croissante.
On pourrait tenter de continuer à étendre la définition de $\mathbb{P}$ (réunion dénombrables d'intervalles, intersections dénombrables de telles unions, etc...). Mais ce n'est pas aisé car la structure des boréliens de $\mathbb{R}$ est compliquée (cf la notion de hiérarchie de Borel).
Le lemme des classes monotones rend cette idée d'extension à une tribu engendrée beaucoup plus limpide, et permet par ailleurs d'étendre ce type de raisonnement à des espaces quelconques.

Définition 1.2. *Soit Ω un espace quelconque, $\mathcal{P}(\Omega)$ l'ensemble de ses parties. La classe $\mathcal{M} \subset \mathcal{P}(\Omega)$ est une classe monotone ssi*

1. $\Omega \in \mathcal{M}$,
2. $(A,B) \in \mathcal{M}^2,\ A \subset B \Rightarrow B \setminus A \in \mathcal{M}$,
3. $(A_n)_{n\geq 0} \in \mathcal{M}^{\mathbb{N}},\ \forall n \geq 0, A_n \subset A_{n+1} \Rightarrow \cup_{n\geq 0} A_n \in \mathcal{M}$.

Il est très facile de démontrer —comme pour les tribus— que l'intersection d'une famille de classes monotones est encore une classe monotone (exercice : vérifier cette assertion), de sorte que, pour une collection quelconque $\mathcal{C}$ de parties de Ω, il existe une plus petite classe monotone contenant $\mathcal{C}$. On la notera $\mathcal{M}(\mathcal{C})$.

Définition 1.3. *On appelle π-système une classe $\mathcal{C} \subset \mathcal{P}(\Omega)$ stable par intersection finie :*

$$A, B \in \mathcal{C} \Rightarrow A \cap B \in \mathcal{C}.$$

On a alors le résultat suivant.

Theorem 1.4 (Lemme des classes monotones)**.** *Si $\mathcal{C}$ est un π-système, alors $\mathcal{M}(\mathcal{C}) = \sigma(\mathcal{C})$.*

On prouve ce lemme un peu plus loin. Commençons tout d'abord par l'appliquer à la preuve de l'unicité de notre mesure $\mathbb{P}$.
Tout d'abord remarquons que la classe

$$\mathcal{C} := \{(-\infty, x], x \in \mathbb{R}\}$$

est bien un π-système, puisque si $x < y$,

$$(-\infty, x] \cap (-\infty, y) = (-\infty, x].$$

Supposons alors que $\mathbb{P}_1, \mathbb{P}_2$ sont deux mesures de probabilité vérifiant (1). On introduit alors

$$\mathcal{M} = \{B \in \mathcal{B}(\mathbb{R}) : \mathbb{P}_1(B) = \mathbb{P}_2(B)\}.$$

Puisque $\mathbb{P}_1, \mathbb{P}_2$ vérifient (1), on a forcément

$$\mathcal{M} \supset \mathcal{C}.$$

Les propriétés de mesure de probabilité, que vérifient $\mathbb{P}_1$ et $\mathbb{P}_2$ impliquent directement que $\mathcal{M}$ est une classe monotone (exercice : vérifier cette assertion). Comme $\mathcal{M}(\mathcal{C})$ est *la plus petite* classe monotone contenant $\mathcal{C}$ on a donc

$$\mathcal{M} \supset \mathcal{M}(\mathcal{C}) \tag{2}$$

D'après le Lemme des classes monotones, on déduit que

$$\mathcal{M} \supset \sigma(\mathcal{C}).$$

Or, tout intervalle ouvert de $\mathbb{R}$ est dans $\sigma(\mathcal{C})$. En effet

$$(-\infty, b) = \bigcup_{n \geq 0} (-\infty, b - 1/n],$$

$$(a, +\infty) = (-\infty, a]^c$$
$$(a, b) = (-\infty, b) \cap (-\infty, a]^c.$$

Quitte à passer aux complémentaires, il est facile de voir que tous les intervalles fermés sont également dans $\sigma(\mathcal{C})$.
Soit désormais un compact K de $\mathbb{R}$. On a (exercice : le vérifier) :

$$K = \bigcap_{n \in \mathbb{N}^*} \bigcup_{x \in K}]x - 1/n, x + 1/n[$$

Cependant, pour tout $n \in \mathbb{N}^*$, $\bigcup_{x \in K}]x - 1/n, x + 1/n[$ est un recouvrement ouvert de K et donc, par Borel-Lebesgue, on peut en extraire un sous-recouvrement fini, i.e. pour tout $n \in \mathbb{N}$ il existe I_n fini et des $x_{i,n}, i \in I_n$ tels que

$$\bigcup_{x \in K}]x - 1/n, x + 1/n[= \bigcup_{i \in I_n}]x_{i,n} - 1/n, x_{i,n} + 1/n[.$$

On en déduit que

$$K = \bigcap_{n \in \mathbb{N}^*} \bigcup_{i \in I_n}]x_{i,n} - 1/n, x_{i,n} + 1/n[\ \in \sigma(\mathcal{C}).$$

Soit désormais un fermé F de $\mathbb{R}$. L'ensemble $F \cap [-k, k]$ est compact, et donc

$$F = \bigcup_{k \in \mathbb{N}} F \cap [-k, k] \in \sigma(\mathcal{C}).$$

On a donc montré que tout fermé de $\mathbb{R}$ est dans $\sigma(\mathcal{C})$. Par passage au complémentaire tout ouvert est donc également dans $\sigma(\mathcal{C})$. Par définition de $\mathcal{B}(\mathbb{R})$, $\sigma(\mathcal{C}) \supset \mathcal{B}(\mathbb{R})$, mais comme $\mathcal{C} \subset \mathcal{B}(\mathbb{R})$ on conclut finalement que

$$\sigma(\mathcal{C}) = \mathcal{B}(\mathbb{R}).$$

D'après (2) on a donc montré que

$$\mathcal{M} = \{B \in \mathcal{B}(\mathbb{R}) : \mathbb{P}_1(B) = \mathbb{P}_2(B)\} = \mathcal{B}(\mathbb{R}),$$

autrement dit, les mesures $\mathbb{P}_1$ et $\mathbb{P}_2$ coiincident sur $\mathcal{B}(\mathbb{R})$, i.e. $\mathbb{P}_1 = \mathbb{P}_2$. $\square$
Pour achever la preuve de l'unicité il nous reste à démontrer le Lemme de classes monotones.
Preuve du Lemme des classes monotones : Commençons par former

$$\mathcal{D}_1 := \{A \in \mathcal{M}(\mathcal{C}) : \forall C \in \mathcal{C}, A \cap C \in \mathcal{M}(\mathcal{C})\}.$$

Il est facile de voir que la classe $\mathcal{D}_1$ est une classe monotone (exercice : le vérifier). Elle contient $\mathcal{C}$ puisque $\mathcal{C}$ est un π-système. Donc elle contient $\mathcal{M}(\mathcal{C})$, la plus petite classe monotone contenant $\mathcal{C}$. Par définition elle est incluse dans $\mathcal{M}(\mathcal{C})$, on conclut que

$$\mathcal{D}_1 = \mathcal{M}(\mathcal{C}).$$

Introduisons alors

$$\mathcal{D}_2 := \{A \in \mathcal{M}(\mathcal{C}) : \forall C \in \mathcal{M}(\mathcal{C}), A \cap C \in \mathcal{M}(\mathcal{C})\}.$$

La classe $\mathcal{D}_1$ est une classe monotone (même argument que pour $\mathcal{D}_1$). Elle contient $\mathcal{C}$ puisque $\mathcal{D}_1 = \mathcal{M}(\mathcal{C})$. Donc elle contient $\mathcal{M}(\mathcal{C})$, la plus petite classe monotone contenant $\mathcal{C}$. Par définition elle est incluse dans $\mathcal{M}(\mathcal{C})$, on conclut que

$$\mathcal{D}_2 = \mathcal{M}(\mathcal{C}).$$

On a donc démontré que $\mathcal{M}(\mathcal{C})$ est un π-système. Et par définition c'est une classe monotone. C'est donc une tribu :

1. $\Omega \in \mathcal{M}(\mathcal{C})$ (première propriété de classe monotone)
2. Si $A \in \mathcal{M}(\mathcal{C})$ alors $\Omega \setminus A = A^c \in \mathcal{M}(\mathcal{C})$ (deuxième propriété de classe monotone)
3. $\mathcal{M}(\mathcal{C})$ est stable par intersections finies (c'est un π-système d'après notre raisonnement précédent). Quitte à passer au complémentaire, $\mathcal{M}(\mathcal{C})$ est donc stable par réunions finies, i.e. si $A_1, ..., A_n \in \mathcal{M}(\mathcal{C})$ alors $\bigcup_{i=1}^{n} A_i \in \mathcal{M}(\mathcal{C})$.

4. Si $A_n \in \mathcal{M}(\mathcal{C})$ pour tout $n \geq 1$, alors les ensembles $B_1 = A_1, B_2 = A_1 \cup A_2, ..., B_n = \bigcup_{i=1}^n A_i, ...$ sont également dans $\mathcal{M}(\mathcal{C})$ d'après le point précédent, et de plus la suite $(B_n)_{n\geq 1}$ est croissante. D'après la troisième propriété de classe monotone on a donc
$$\bigcup_{n\geq 1} A_n = \bigcup_{n\geq 1} B_n \in \mathcal{M}(\mathcal{C}).$$

Ainsi $\mathcal{M}(\mathcal{C})$ est une tribu qui contient $\mathcal{C}$, la plus petite tribu contenant $\mathcal{C}$. On a donc
$$\mathcal{M}(\mathcal{C}) \supset \sigma(\mathcal{C}).$$
Par ailleurs la tribu $\sigma(\mathcal{C})$ est évidemment une classe monotone, et comme $\mathcal{M}(\mathcal{C})$ est la plus petite classe monotone contenant $\mathcal{C}$ on a
$$\mathcal{M}(\mathcal{C}) \subset \sigma(\mathcal{C}).$$
On conclut finalement
$$\mathcal{M}(\mathcal{C}) = \sigma(\mathcal{C}). \qquad \square$$

1.2 Existence

On a vu dans la preuve de l'unicité qu'il était direct d'étendre $\mathbb{P}$ à la classe des intervalles de $\mathbb{R}$ (c'est d'ailleurs en faisant ces extensions qu'on a utilisé les hypothèses sur la fonction F).
Pour $B \in \mathcal{B}(\mathbb{R})$ on définit alors
$$\mathbb{P}^*(B) = \inf_{\{(A_n)_{n\geq 1} \text{ intervalles de } \mathbb{R}\ :\ \bigcup_{n\geq 1} A_n \supset B\}} \sum_{n\geq 1} \mathbb{P}(A_n),$$

Lemme 1.5. $\mathbb{P}^*$ *est une* mesure extérieure*, c'est-à-dire que*

1. $\mathbb{P}^*(\emptyset) = 0$.

2. si $B_1 \subset B_2$ *sont des boréliens,* $\mathbb{P}^*(B_1) \leq \mathbb{P}^*(B_2)$.

3. si les $B_n, n \geq 1$ *sont des boréliens,* $\mathbb{P}^*(\bigcup_{n\geq 1} B_n) \leq \sum_{n\geq 1} \mathbb{P}^*(B_n)$.

Preuve : Pour la première propriété ci-dessus, on peut utiliser par exemple que $\emptyset \subset (-\infty, x]$ pour tout $x \in \mathbb{R}$, et donc $\mathbb{P}^*(\emptyset) \leq F(x)$ pour tout $x \in \mathbb{R}$. On conclut grâce au fait que $\lim_{x\to-\infty} F(x) = 0$.
Pour la deuxième il suffit de voir que si $(A_n)_{n\geq 1}$ est une suite d'intervalles de $\mathbb{R}$ telle que $\bigcup_{n\geq 1} A_n \supset B_2$ alors forcément cette union contient également $B_1 \subset B_2$.
Pour la dernière propriété, on peut supposer de plus que $\mathbb{P}^*(B_n) < \infty$ pour tout $n \geq 1$ (sinon l'inégalité est évidente). Fixons alors $\varepsilon > 0$. Par définition de $\mathbb{P}^*$, on a, pour tout $n \geq 1$, l'existence d'une suite d'intervalles $(A_{n,k})_{k\geq 1}$ telle que $B_n \subset \bigcup_{k\geq 1} A_{n,k}$ et
$$\sum_{k\geq 1} \mathbb{P}(A_{n,k}) \leq \mathbb{P}^*(B_n) + \varepsilon 2^{-n}. \tag{3}$$

Il faut alors observer que

$$\bigcup_{n\geq 1} B_n \subset \bigcup_{n\geq 1}\bigcup_{k\geq 1} A_{n,k},$$

et donc par définition de $\mathbb{P}^*$,

$$\mathbb{P}^*(\bigcup_{n\geq 1} B_n) \leq \sum_{n\geq 1}\sum_{k\geq 1} \mathbb{P}(A_{n,k})$$

D'après (3) on a donc

$$\mathbb{P}^*(\bigcup_{n\geq 1} B_n) \leq \sum_{n\geq 1} \mathbb{P}^*(B_n) + \varepsilon,$$

et on conclut puisque ε est arbitraire. $\square$

Lemme 1.6. *La classe*

$$\mathcal{G} := \{G \in \mathcal{B}(\mathbb{R}) \ : \ \forall B \in \mathcal{B}(\mathbb{R}) \ \ \mathbb{P}^*(B) = \mathbb{P}^*(B\cap G) + \mathbb{P}^*(B \cap G^c)$$

est une tribu.
Par ailleurs si $(G_n)_{n\geq 1}$ est une suite d'éléments disjoints de $\mathcal{G}$ on a

$$\mathbb{P}^*(\bigcup_{n\geq 1} G_n) = \sum_{n\geq 1} \mathbb{P}^*(G_n).$$

Preuve :

1. Il est évident que $\Omega \in \mathcal{G}$ puisque $\mathbb{P}^*(\emptyset) = 0$.
2. Si $G \in \mathcal{G}$ alors pour tout $B \in \mathcal{B}(\mathbb{R})$,

$$\mathbb{P}^*(B) = \mathbb{P}^*(B\cap G) + \mathbb{P}^*(B\cap G^c) = \mathbb{P}^*(B \cap (G^c)^c) + \mathbb{P}^*(B\cap G^c),$$

   et donc $G^c \in \mathcal{G}$
3. Si $G_1 \in \mathcal{G}, G_2 \in \mathcal{G}$, notons $G = G_1 \cap G_2$ et montrons que $G \in \mathcal{G}$.
   Observons que $G^c \cap G_2 = G_2 \cap G_1^c$, tandis que $G^c \cap G_2^c = G_2^c$. Or $G_2 \in \mathcal{G}$ donc pour $B \in \mathcal{B}(\mathbb{R})$,

$$\mathbb{P}^*(G^c \cap B) = \mathbb{P}^*(G_2 \cap G_1^c \cap B) + \mathbb{P}^*(G_2^c \cap B)$$

   Par ailleurs, toujours puisque $G_2 \in \mathcal{G}$,

$$\mathbb{P}^*(B) = \mathbb{P}^*(G_2^c \cap B) + \mathbb{P}^*(G_2 \cap B).$$

   Enfin, puisque $G_1 \in \mathcal{G}$,

$$\mathbb{P}^*(G_2 \cap B) = \mathbb{P}^*(G \cap B) + \mathbb{P}^*(G_2 \cap G_1^c \cap B)$$

   En combinant les trois égalités, on obtient finalement

$$\begin{aligned} & \mathbb{P}^*(G\cap B) + \mathbb{P}^*(G^c \cap B) \\ = \ & \mathbb{P}^*(G_2 \cap B) - \mathbb{P}^*(G_2 \cap G_1^c \cap B) + \mathbb{P}^*(G_2 \cap G_1^c \cap B) + \mathbb{P}^*(G_2^c \cap B) = \mathbb{P}^*(B). \end{aligned}$$

4. Si $G_1 \in \mathcal{G}$, $G_2 \in \mathcal{G}$ sont disjoints, notons $G = G_1 \cup G_2$ et montrons que $G \in \mathcal{G}$. On a
$$G \cap G_1 = G_1, \quad G \cap G_1^c = G_2.$$
Puisque $G_1 \in \mathcal{G}$ il vient, comme souhaité,
$$\mathbb{P}^*(G \cap B) = \mathbb{P}^*(G_1 \cap B) + \mathbb{P}^*(G_2 \cap B).$$
Notons en particulier que
$$\mathbb{P}^*(G) = \mathbb{P}^*(G_1) + \mathbb{P}^*(G_2),$$
et que notre argument se généralise immédiatemment à des réunions disjointes, finies.
5. Si les $G_n, n \geq 1$ sont des éléments de $\mathcal{G}$ disjoints, notons $G = \bigcup_{n\geq 1} G_n$ et montrons [1] que $G \in \mathcal{G}$. D'après le point précédent on a, pour tout $n \geq 1$, $H_n := \bigcup_{i=1}^n G_i \in \mathcal{G}$. Observons par ailleurs que $H_n^c \subset G^c$. On a donc pour tout $B \in \mathcal{B}(\mathbb{R})$, en utilisant le point 3.,
$$\begin{aligned} \mathbb{P}^*(B) &= \mathbb{P}^*(H_n \cap B) + \mathbb{P}^*(H_n^c \cap B) \\ &= \sum_{i=1}^n \mathbb{P}^*(G_i \cap B) + \mathbb{P}^*(H_n^c \cap B) \\ &\geq \sum_{i=1}^n \mathbb{P}^*(G_i \cap B) + \mathbb{P}^*(G^c \cap B). \end{aligned}$$
Comme le raisonnement est valable pour tout $n \geq 1$,
$$\mathbb{P}^*(B) \geq \sum_{n\geq 1} \mathbb{P}^*(G_n \cap B) + \mathbb{P}^*(G^c \cap B). \tag{4}$$
En utilisant la troisième propriété de mesure extérieure, on a donc
$$\mathbb{P}^*(B) \geq \mathbb{P}^*(G \cap B) + \mathbb{P}^*(G^c \cap B).$$
Or cette troisième propriété, cette fois pour $B = (B \cap G) \cup (B \cap G^c)$ fournit l'inégalité inverse et on conclut comme souhaité que
$$\mathbb{P}^*(B) = \mathbb{P}^*(G \cap B) + \mathbb{P}^*(G^c \cap B).$$
6. Supposons à nouveau que les $G_n, n \geq 1$ sont des éléments de $\mathcal{G}$ disjoints, $H_n = \bigcup_{k=1}^n G_k, G = \bigcup_{n\geq 1} G$. Quitte à prendre $B = G$ dans (4), on a
$$\mathbb{P}^*(G) \geq \sum_{n\geq 1} \mathbb{P}^*(G_n).$$

1. Si les $G_n, n \geq 1$ sont des éléments de $\mathcal{G}$ quelconques, il suffit de former la suite d'éléments de $\mathcal{G}$ disjoints $H_1 = G_1, H_2 = G_2 \cap G_1^c, ..., H_n = G_n \cap (\bigcup_{k\leq n-1} G_k)^c, ...$, et observer que $\bigcup_{n\geq 1} H_n = \bigcup_{n\geq 1} G_n$.

D'après la troisième propriété de mesure extérieure, on a l'inégalité inverse, et on conclut que

$$\mathbb{P}^*(G) = \sum_{n\geq 1} \mathbb{P}^*(G_n). \qquad \square$$

Les Lemmes 1.5, 1.6 permettent de conclure que $\mathbb{P}^*$ est une mesure sur $\mathcal{G}$ (c'est même une mesure de probabilité puisque $\mathbb{P}^*(\mathbb{R}) = \lim_{x\to\infty} F(x) = 1$).
Reste à observer que $\mathcal{G}$ contient forcément $\mathcal{C}$. Un $C \in \mathcal{C}$ s'écrit forcément $(-\infty, x]$ pour un x réél. Fixons alors $B \in \mathcal{B}(\mathbb{R})$. On sait que

$$\mathbb{P}^*(B) = \inf_{\{(A_n)_{n\geq 1} \text{ intervalles de } \mathbb{R} \,:\, \bigcup_{n\geq 1} A_n \supset B\}} \sum_{n\geq 1} \mathbb{P}(A_n).$$

Pour $\varepsilon > 0$ fixé, on peut donc trouver des intervalles $(A_n)_{n\geq 1}$ tels que

$$\sum_{n\geq 1} \mathbb{P}(A_n) \leq \mathbb{P}^*(B) + \varepsilon.$$

Formons alors la suite d'intervalles

$$C_{2n} = A_n \cap (-\infty, x], C_{2n+1} = A_n \cap (x, +\infty),$$

de sorte que $\mathbb{P}(A_n) = \mathbb{P}(C_{2n}) + \mathbb{P}(C_{2n+1})$, et

$$\bigcup_{n\geq 1} C_{2n} \supset B \cap C, \quad \bigcup_{n\geq 1} C_{2n} \supset B \cap C^c$$

On a donc pour tout $\varepsilon > 0$,

$$\mathbb{P}^*(B \cap C) + \mathbb{P}^*(B \cap C^c) \leq \mathbb{P}^*(B) + \varepsilon.$$

Comme $\varepsilon > 0$ est arbitraire, on conclut que $\mathbb{P}^*(B \cap C) + \mathbb{P}^*(B \cap C^c) \leq \mathbb{P}^*(B)$.
L'inégalité $\mathbb{P}^*(B\cap C) + \mathbb{P}^*(B\cap C^c) \geq \mathbb{P}^*(B)$ découle simplement des propriétés de mesure extérieure. Comme le raisonnement est valable quelque soit B, on déduit que $C \in \mathcal{G}$.
On a donc que $\mathcal{C} \subset \mathcal{G}$, et comme $\mathcal{G}$ est une tribu, $\mathcal{G} \supset \sigma(\mathcal{C})$. Or on a déjà montré (cf la fin de la preuve d'unicité) que $\sigma(\mathcal{C}) = \mathcal{B}(\mathbb{R})$, et donc $\mathcal{G} = \mathcal{B}(\mathbb{R})$.
On conclut donc, comme souhaité, que $\mathbb{P}^*$ est une mesure de probabilité sur $\mathcal{B}(\mathbb{R})$. $\square$

1.3 Le cas de v.a. à valeurs dans E

. On a traité le cas $E = \mathbb{R}$. En fait, on peut faire des raisonnement quasi-identiques dans les cas de $E = \mathbb{R}^n$ (ce qui permet de comprendre comment caractériser la loi jointe d'un n-uplet de variables aléatoires réelles), $E = \mathbb{R}^{\mathbb{N}}$ (ce qui permet, via le théorème d'extension de Kolmogorov, de comprendre comment caractériser la loi d'une suite de v.a.r., i.e. un processus discret).
Certains des résultats utilisés dans les paragraphes précédents sont en fait très généraux :

a1. Le lemme des classes monotones concerne un espace Ω quelconque.

a2. Dans la preuve d'unicité, on fait justement appel aux lemme des classes monotones. L'idée de cette preuve est tout à fait générale. Précisément, si $\mathcal{C}$ est un π-système, et si $\mathbb{P}_1$ et $\mathbb{P}_2$ sont deux mesures qui coıincident sur $\mathcal{C}$, alors

$$\mathcal{M} = \{A \in \sigma(\mathcal{C}) : \mathbb{P}_1(A) = \mathbb{P}_2(A)\}$$

est une classe monotone, qui est donc égale à $\sigma(\mathcal{C})$.

On a donc dans ce cas que $\mathbb{P}_1$ et $\mathbb{P}_2$ coıincident sur $\sigma(\mathcal{C})$.

Ainsi, lorsqu'on veut caractériser une mesure sur une tribu $\mathcal{B}$, il suffit de le faire sur un π-système qui engendre $\mathcal{B}$.

Par exemple, pour caractériser une mesure sur $\mathbb{R}^d$ il suffit en fait de se la donner sur les pavés ouverts.

Pour caractériser une mesure sur $\mathbb{R}^{\mathbb{N}}$, il suffit de se donner des mesures $\mathbb{P}_n$ sur $\mathbb{R}^n, n \geq 0$, pourvu qu'elles vérifient une propriété de consistance très naturelle ; précisément, la restriction de $\mathbb{P}_{n+k}$, mesure sur $\mathbb{R}^{n+k}$, aux événements ne dépendant que des n premières coordonnées, se doit forcément de coıincider avec $\mathbb{P}_n$.

b. Dans la preuve de l'existence, on a utilisé la notion de mesure extérieure, qui est tout à fait générale. La preuve que $\mathcal{G}$ est une tribu est également générale.

En revanche, pour montrer que $\mathcal{C} \subset \mathcal{G}$ on s'est servi des propriétés de $\mathbb{R}$.

Il est cependant possible de généraliser pourvu que l'espace E soit métrique et séparable. C'est bien sûr le cas de $E = \mathbb{R}^d, E = \mathbb{R}^{\mathbb{N}}$.