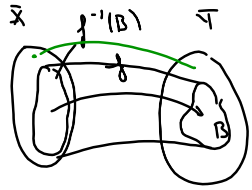
# Distributions

**Background:** If $f: X \to Y$ is a function and $B$ is a subset of $Y$, then the inverse image of $B$ is

$$f^{-1}(B) = \{x \in X : f(x) \in B\}$$




**Proposition:**

(i) $f^{-1}(\emptyset) = \emptyset$ and $f^{-1}(Y) = X$

(ii) $f^{-1}(\underbrace{B^c}_{\text{complement wrt } Y}) = \underbrace{(f^{-1}(B))}_{\text{compl. wrt to } X}{}^c$

(iii) If $\{A_i\}_{i \in I}$ is a family of sets, then

$$f^{-1}\Big(\bigcup_{i \in I} A_i\Big) = \bigcup_{i \in I} f^{-1}(A_i) \quad \text{and} \quad f^{-1}\Big(\bigcap_{i \in I} A_i\Big) = \bigcap_{i \in I} f^{-1}(A_i).$$

---

Distribution function of $X$: $F(x) = \underbrace{P\{\omega : X(\omega) \le x\}}_{P\{X \le x\}}$

Want to extend this notion to a measure called the distribution of $X$.

**Proposition:** Assume that $(\Omega, \mathcal{F}, P)$ is a prob. space and that $X : \Omega \to \mathbb{R}$ is a random variable. Then

(*) $X^{-1}(B) = \{\omega : X(\omega) \in B\} \in \mathcal{F}$

for all Borel sets $B$.

**Proof:** Define

$$\mathcal{A} = \{B \subseteq \mathbb{R} : X^{-1}(B) \in \mathcal{F}\}$$

We want to prove that $\mathcal{A}$ is a $\sigma$-algebra containing all half-open intervals $(a,b]$. Since the Borel $\sigma$-algebra $\mathcal{B}$ is the smallest $\sigma$-algebra containing the half-open intervals, this will tell us that $\mathcal{B} \subseteq \mathcal{A}$, and hence (*) is satisfied by all $B \in \mathcal{B}$.

Need to check that $\mathcal{A}$ satisfies the conditions of a $\sigma$-algebra containing all half-open intervals.

First note that

$$X^{-1}((a,b]) = \overbrace{\{\omega : X(\omega) \le b\}}^{\in \mathcal{F}} \cap \overbrace{\{\omega : X(\omega) > a\}}^{\in \mathcal{F}} \in \mathcal{F}.$$

Check the conditions for a $\sigma$-algebra.

$\mathcal{A} = \{B \subseteq \mathbb{R} : X^{-1}(B) \in \mathcal{F}\}$

(i) $\emptyset \in \mathcal{A}$: $X^{-1}(\emptyset) = \emptyset \in \mathcal{F}$.

(ii) Assume that $B \in \mathcal{A}$, need show that $B^c \in \mathcal{A}$. But $X^{-1}(B^c) = (\underbrace{X^{-1}(B)}_{\in \mathcal{F}})^c \in \mathcal{F}$.

(iii) Assume that $A_n \in \mathcal{A}$ for all $n$. Need to prove that $\bigcup_{n \in \mathbb{N}} A_n \in \mathcal{A}$. But

$$X^{-1}\Big(\bigcup_{n \in \mathbb{N}} A_n\Big) = \bigcup_{n \in \mathbb{N}} \underbrace{X^{-1}(A_n)}_{\in \mathcal{F}} \in \mathcal{F}.$$

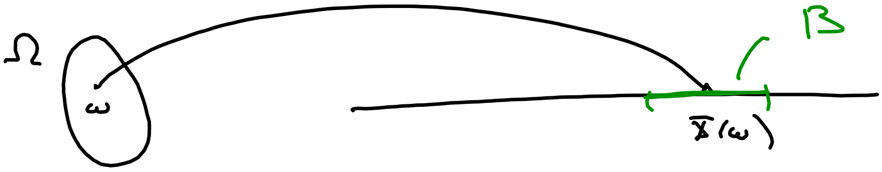
Definition: Assume that $X$ is a random variable and $\mathcal{B}$ is the Borel $\sigma$-algebra on $\mathbb{R}$. Define

$$\mu : \mathcal{B} \to \mathbb{R}$$

by

$$\mu(B) = P(\underbrace{X^{-1}(B)}_{\mathcal{F}})$$

We call $\mu$ the distribution of $X$.




The probability that $X(\omega)$ is in $B$:

$$P\{\omega : X(\omega) \in B\} = P\{X^{-1}(B)\} = \mu(B)$$

Theorem: If $X$ is a random variable, its distribution $\mu$ is a probability measure on $(\mathbb{R}, \mathcal{B})$.

Proof: Check the conditions for a prob. measure:

$$\boxed{\mu(B) = P\{X^{-1}(B)\}}$$

(i) $\mu(\phi) = P\{X^{-1}(\phi)\} = P(\phi) = 0$.

(ii) $\mu(\mathbb{R}) = P\{\underbrace{X^{-1}(\mathbb{R})}_{\Omega}\} = P(\Omega) = 1$.

(iii) If $\{B_n\}$ is a disjoint sequence,

$$\mu(\bigcup_{n \in \mathbb{N}} B_n) = P\{X^{-1}(\bigcup_{n\in\mathbb{N}} B_n)\} = P\{\bigcup_{n \in \mathbb{N}} \overbrace{X^{-1}(B_n)}^{\text{disjoint}}\}$$

$$= \sum_{n \in \mathbb{N}} P[X^{-1}(B_n)] = \sum_{n \in \mathbb{N}} \mu(B_n)$$

---

Two descriptions of the distribution of a random variable $X$.

Distribution function $F : \mathbb{R} \to [0,1]$ : $F(x) = P\{\omega : X(\omega) \le x\} = P(X^{-1}(-\infty, x])$

Distribution: Prob. measure on $(\mathbb{R}, \mathcal{B})$ : $\mu(B) = P\{X^{-1}(B)\}$

Relationship: $F(x) = P(X^{-1}(-\infty, x]) = \mu((-\infty, x])$

# Distribution function

Definition: A distribution function is a function $F: \mathbb{R} \to \mathbb{R}$ such that

(i) $F$ is a increasing and right continuous function.

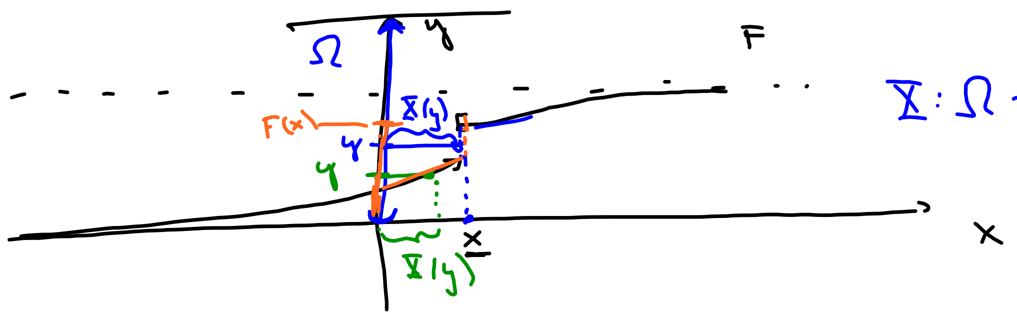
(ii) $\lim_{x \to \infty} F(x) = 1$ and $\lim_{x \to -\infty} F(x) = 0$.

Theorem: A distribution function is always the distribution function of a random variable.

Sketch of proof: Assume $F$ is a distribution. I need a prob. space and a random variable with $F$ as its distribution.

$\Omega = (0,1)$, $\mathcal{B}$ = Borel σ-algebra on $(0,1)$, $P$ = the Lebesgue measure on $\mathcal{B}$

$P(I)$ = length of $I$ $\quad ( P((a,b]) = b - a )$




$X: \Omega \to \mathbb{R}$

Formal definition: $X(y) = \inf \{ x \in \mathbb{R} : F(x) \geq y \}$

Need to show that $F$ is the distribution function of $X$

$$P\{ y \in (0,1) : X(y) \leq x \} = \underbrace{P\{ (0, F(x)] \}}_{\text{length of interval}} = F(x)$$

and hence the distribution function of $X$ is $F$.

Independence of random variables

**Definition:** Two random variables $X, Y$ are independent if

$$P[\underbrace{X \le x}_{X(\omega)\in(-\infty,x]} \text{ and } \underbrace{Y \le y}_{Y(\omega)\in(-\infty,y]}] = P[X \le x]P[Y \le y] \text{ for all } x,y \in \mathbb{R}.$$

(this just means that the sets $\{X \le x\}, \{Y \le y\}$ are independent for all $x, y$)

**Definition:** A family $\{X_i\}_{i\in I}$ of random variables is independent if whenever we choose $\overbrace{i_1, i_2, .., i_n}^{\text{distinct}} \in I$ and $x_1, x_2, .., x_n \in \mathbb{R}$, then

$$P[X_{i_1} \le x_1 \text{ and } X_{i_2} \le x_2 \text{ and} .. \text{ and } X_{i_n} \le x_n]$$
$$= P[X_{i_1} \le x_1]P[X_{i_2} \le x_2] \dots P[X_{i_n} \le x_n]$$

**Theorem:** Two random variables $X$ and $Y$ are independent iff

(*) $P[X \in A \text{ and } Y \in B] = P[X \in A]P[Y \in B]$ for all Borel sets $A, B$.

**Proof:** If (*) holds, then $X$ and $Y$ are independent (just choose $A = (-\infty, x]$ and $B = (-\infty, y]$).

The hard part is to prove that if $X, Y$ are independent, then (*) holds.

**Plan:** Three steps:

$\mathcal{A}_0$ = the algebra generated by the half-open intervals
= the collection of all finite disjoint unions of half-open intervals (problem last week)

**Step I:** Prove that

$$P[X \in A \wedge Y \le y] = P[X \in A]P[Y \le y] \text{ for } A \in \mathcal{A}_0$$

**Step II:** Define

$$\mathcal{M}_1 = \{A \in \mathcal{F} : P[X \in A \wedge Y \le y] = P[X \in A]P[Y \le y] \text{ for } y \in \mathbb{R}\}$$

By step I, $\mathcal{A}_0 \subseteq \mathcal{M}_1$. We are going to prove that $\mathcal{M}_1$ is a monotone class. By the MCT, then

$$\mathcal{M}_1 \supseteq \mathcal{M}(\mathcal{A}_0) = \sigma(\mathcal{A}_0) = \mathcal{B}.$$

Hence $P[X \in A \wedge Y \le y] = P[X \in A]P[Y \le y]$ hold for all $Y \in (-\infty, y]$, $A \in \mathcal{B}$.

**Step 3:** Define

$$\mathcal{M}_2 = \{B \in \mathcal{F} : P[X \in A \wedge Y \in B] = P[X \in A]P[Y \in B] \text{ for all } A \in \mathcal{B}\}$$

Want to prove $\mathcal{A}_0 \subseteq \mathcal{M}_2$ and that $\mathcal{M}_2$ is a monotone class.
If so, then

$$\mathcal{M}_2 \supseteq \mathcal{M}(\mathcal{A}_0) = \sigma(\mathcal{A}_0) = \mathcal{B}.$$

This means that if $B \in \mathcal{B} \subseteq \mathcal{M}_2$, then

$$P[X \in A \wedge Y \in B] = P[X \in A]P[Y \in B] \text{ for all } A \in \mathcal{B}. \quad \text{QED.}$$

A typical class argument: $\mathcal{M}_1 = \{A \in \mathcal{F} : P[X \in A \wedge Y \le y] = P[X \in A]P[Y \le y] \text{ for } y \in \mathbb{R}\}$

$\{A_n\}$ increasing and in $\mathcal{M}_1$

$$P[X \in \bigcup_{n\in\mathbb{N}} A_n \wedge Y \le y] \overset{\text{cont of meas}}{=} \lim_{n\to\infty} P[X \in A_n \wedge Y \le y]$$

$$= \lim_{n\to\infty} P[X \in A_n]P[Y \le y] \overset{\text{cont of meas}}{=} P[X \in \bigcup A_n]P[Y \le y]$$