Inverse Map vs Inverse Function

A Short Note on a Distinction That Matters

1. The Core Distinction

The notation f^{-1} is overloaded in mathematics. It denotes two entirely different objects depending on context, and conflating them leads to genuine errors in topology, measure theory, and probability.

Definition — Inverse Function

Given $f : X \to Y$, the **inverse function** $f^{-1} : Y \to X$ sends each $y \in Y$ to the unique $x \in X$ with $f(x) = y$. It exists **if and only if** f is bijective (injective and surjective).

Definition — Inverse Map (Pre-image)

Given **any** function $f : X \to Y$ (no conditions on f), the **inverse map** $f^{-1} : P(Y) \to P(X)$ acts on subsets:

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

This is a map between power sets. It **always exists**, regardless of whether f is injective, surjective, or neither.

2. Examples

Example 1 — A non-injective function

Let $f : R \to R$, $f(x) = x^2$.

Inverse function? Does not exist. f is not injective ($f(2) = f(-2) = 4$) and not surjective onto R (negative numbers have no pre-image).

Inverse map? Always works, on subsets:

$$f^{-1}(\{4\}) = \{-2, 2\} \quad f^{-1}(\{0, 9\}) = \{-3, 0, 3\} \quad f^{-1}((-\infty, 0)) = \varnothing \quad f^{-1}([0, 1]) = [-1, 1]$$

Every subset of R has a well-defined pre-image, even though no element of R has a well-defined "inverse value."

Example 2 — Where both exist but differ

Let $f : R \to R$, $f(x) = 2x + 1$ (bijective).

Inverse function: $f^{-1}(y) = (y - 1)/2$, so $f^{-1}(5) = 2$. This is element-to-element.

Inverse map: $f^{-1}(\{5\}) = \{2\}$, $f^{-1}([1, 3]) = [0, 1]$. This is subset-to-subset.

When f is bijective, the inverse map applied to a singleton {y} gives the singleton {x}, and the two notions are consistent — but they live in different mathematical worlds.

Example 3 — Topology

Let $f : R \to R$, $f(x) = \sin x$. No inverse function exists. But the entire definition of topological continuity is:

$$\text{f is continuous ■ } f^{-1}(U) \text{ is open for every open } U \subseteq R$$

This uses only the inverse map. The inverse function is irrelevant.

3. Summary Comparison

	Inverse Function	Inverse Map (Pre-image)
Domain	$Y \to X$ (elements)	$P(Y) \to P(X)$ (subsets)
Exists when	f is bijective only	Always, for any f
f■¹({y}) gives	A single element x	A subset of X (possibly ∅)
Key property	f ■ f■¹ = id	Preserves ∪, ∩, complements
Lives in	Algebra, calculus	Topology, measure, probability

4. The Probability Bridge

A random variable is a measurable function $X : (\Omega, F, P) \to (R, B)$. The word "measurable" means precisely one thing:

$$\text{For every Borel set } B \in B,\ X^{-1}(B) \in F$$

This is the inverse map, not the inverse function. X need not be invertible — many outcomes can map to the same real number. What matters is that every time you ask "what is the probability that X falls in B?", the pre-image $X^{-1}(B)$ lands in the σ-algebra F where P is defined. Only then can you write:

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

Key Insight

P only knows how to measure sets in F. When someone asks "what is the probability that X lands in some Borel set B?", the inverse map is the **only mechanism** that translates that question from R back to Ω where P lives. That translation is the entire content of the word "measurable."

Further consequences built on the inverse map:

Induced measure (distribution). $P_X(B) = P(X^{-1}(B))$ for every Borel set B. The entire transfer from the abstract probability space to the real line is the inverse map.

σ-algebra generated by X. $\sigma(X) = \{X^{-1}(B) : B \in B\}$. This is exactly the "information" that X reveals.

Conditional expectation. E[Y | X] means "the best prediction of Y using only information in σ(X)." Built entirely on the inverse map.

5. Why the Forward Image Cannot Replace It

Critical Asymmetry

The pre-image map preserves arbitrary unions ($f^{-1}(\cup B_\alpha) = \cup f^{-1}(B_\alpha)$), arbitrary intersections, **and** complements. The forward image f(A) does **not** preserve intersections or complements. This is precisely why continuity and measurability are defined through pre-images, never through forward images.

Counterexample: Let $f : R \to R$, $f(x) = x^2$, A = [–2, 0], B = [0, 2]. Then $A \cap B = \{0\}$ and $f(A \cap B) = \{0\}$. But f(A) = [0, 4] and f(B) = [0, 4], so $f(A) \cap f(B) = [0, 4] \neq \{0\}$. The forward image breaks intersections.

6. Simmons' Treatment in *Introduction to Topology and Modern Analysis*

Simmons develops the set-theoretic machinery of the inverse map early and deliberately, in Chapter 1 ("Sets and Functions"), before any topology is introduced. This placement is not accidental — it signals that the pre-image is infrastructure, not a side remark.

Section 2 — Functions (paraphrased key results)

Simmons defines a function $f : X \to Y$ and immediately introduces the forward image $f(A) = \{f(x) : x \in A\}$ and the inverse image $f^{-1}(B) = \{x : f(x) \in B\}$ as operations on subsets. He is careful to note that the notation $f^{-1}(B)$ is defined for any function f, with no assumption of bijectivity.

Key Results Established (paraphrased from Simmons, Ch. 1, §2)

For any function $f : X \to Y$, subsets $A, A_i \subseteq X$, and subsets $B, B_i \subseteq Y$:

The inverse map preserves all set operations:

(a) $f^{-1}(\cup B_i) = \cup f^{-1}(B_i)$ — preserves arbitrary unions

(b) $f^{-1}(\cap B_i) = \cap f^{-1}(B_i)$ — preserves arbitrary intersections

(c) $f^{-1}(B^c) = (f^{-1}(B))^c$ — preserves complements

The forward image is weaker:

(d) $f(\cup A_i) = \cup f(A_i)$ — preserves unions (this one works)

(e) $f(\cap A_i) \subseteq \cap f(A_i)$ — only inclusion, not equality!

(f) $f(A^c)$ has no clean relationship to $f(A)^c$

Simmons notes that equality in (e) holds if and only if f is injective. This asymmetry — the inverse map being a perfect set-operation homomorphism while the forward image is not — is the structural reason why topology, measure theory, and probability are all built on pre-images.

How Simmons uses this later

In Chapter 3, when Simmons defines topological continuity, the definition reads: f is continuous if and only if $f^{-1}(G)$ is open for every open set G. He does not need to check whether f is invertible as a function — the pre-image map is already in his toolkit from Chapter 1. The same machinery reappears in his treatment of open maps, closed maps, and homeomorphisms throughout the book.

Simmons also proves in the problems of Chapter 1 that $f^{-1}(f(A)) \supseteq A$ with equality iff f is injective, and $f(f^{-1}(B)) \subseteq B$ with equality iff f is surjective. These "round-trip" inclusions further clarify why the inverse map and the inverse function are different objects — the round-trip is exact only when f is bijective, i.e., only when the inverse function also exists.

Pedagogical Remark

Simmons' choice to develop the inverse map as pure set theory before introducing any topology is a masterclass in mathematical exposition. By the time the reader reaches the definition of continuity, the pre-image is already a familiar friend — not a new abstraction introduced mid-argument. This is the Simmons method: build the infrastructure quietly, then let the theorems arrive as natural consequences.

7. References

[1] G. F. Simmons, *Introduction to Topology and Modern Analysis*, McGraw-Hill, 1963. Chapter 1, §2 (Functions), §6 (Products and partitions). Chapter 3 (Topological Spaces), §15 (Continuous functions).

[2] P. R. Halmos, *Measure Theory*, Van Nostrand, 1950. Chapter II, §5 (Measurable functions). The pre-image as the defining mechanism for measurability.

[3] P. Billingsley, *Probability and Measure*, 3rd ed., Wiley, 1995. Chapter 1 (Probability spaces and random variables). The induced measure $P_X(B) = P(X^{-1}(B))$.

[4] J. R. Munkres, *Topology*, 2nd ed., Prentice Hall, 2000. Chapter 1, §2 (Functions). Parallel treatment of forward image vs pre-image with the same asymmetry Simmons identifies.

"The distinction between a function and a set-valued mapping is one of the quiet foundations on which analysis is built."

— In the spirit of G. F. Simmons