Probability

heng

2021.10.9

Contents

1 Condition and Independence

These are what distinguish probability from measure theory.

Definition 1 (Conditional probability)**.** *Given a probability space* $(\Omega, \mathcal{F}, \mathbb{P})$*, a random variable* X*, an event* E*, define a finite measure* m *on* Ω *as*

$$m(E') := P([X \in E'] \cap E)$$

This measure is absolutely continuous with respect to the distribution of X *(denote it as* P_X*) since*

$$m(E') = P([X \in E'] \cap E) \leq P([X \in E']) = P_X(E')$$

Therefore, by Radon-Nikodym theorem, there exists a P_X*-integrable function* ψ *s.t.*

$$P([X \in E'] \cap E) = \int_{E'} \psi(\omega) dP_X(\omega)$$

We define $P(E'|X = \omega) := \psi(\omega)$ *as the probability of* E' *under the condition* $X = \omega$*.*

2 Convergence of Random Variables

Definition 2. *Let* $\{f_n\}, f$ *be measurable functions,* μ *be a measure, define*

- ***almost everywhere convergence****: Converge point-wise outside a set with measure 0.*

- ***convergence by measure:*** $\forall \varepsilon > 0, \mu(|f_n - f| > \varepsilon) \to 0$
- ***convergence by distribution:*** $\forall g \in C_{bounded}(\mathbb{R}), \int_{\mathbb{R}} gdf_n \to \int_{\mathbb{R}} gdf$
-

3 Law of Large Numbers

Definition 3. *Let $\{X_n\}$ be r.v., it obeys the* ***Law of large numbers*** *if $\exists \{a_n\}, 0 \leq \{b_n\} \uparrow \infty$ s.t.*

$$P(|\frac{\sum_{i=1}^{n} x_i - a_n}{b_n}| \geq \varepsilon) \to 0$$

for any $\varepsilon > 0$.

Theorem 1. ***Bernoulli's Weak Law of Large Numbers*** *Let μ_n be Bernoullian r.v. ,then $P(|\frac{\mu_n}{n} - p| \geq \varepsilon) \to 0$*

Proof. By applying Chebyshev's inequal., we have

$$P(|\frac{\mu_n}{n} - p| \geq \varepsilon) \leq \frac{1}{n^2\varepsilon^2} D\mu_n = \frac{1}{n^2\varepsilon^2} np(1-p) \to 0$$

□

Example 1. ***Bernstein's proof of Weierstrass approximation theorem***
Let $f \in C^0([0,1])$,

$$B_n(x) := \sum_{i=0}^{n} f(\frac{i}{n}) C_n^i x^i (1-x)^{n-i}$$

then $B_n \rightrightarrows f$.

Proof. Let $\mathcal{X}_i \sim b(1,x)$, we have

$$Ef(\frac{\sum_{i=1}^{n} \mathcal{X}_i}{n}) = \sum_{i=1}^{n} f(\frac{i}{n}) C_n^i x^i (1-x)^{n-i} = B_n(x)$$

therefore,

$$\begin{aligned}
&|Ef(\frac{\sum_{i=1}^{n} \mathcal{X}_i}{n}) - f(x)| \\
\leq &E|f(\frac{\sum_{i=1}^{n} \mathcal{X}_i}{n}) - f(x)| I_{|\frac{\sum_{i=1}^{n} \mathcal{X}_i}{n} - x| > \delta} \\
+ &E|f(\frac{\sum_{i=1}^{n} \mathcal{X}_i}{n}) - f(x)| I_{|\frac{\sum_{i=1}^{n} \mathcal{X}_i}{n} - x| \leq \delta} \\
\leq &2||f||_\infty P(\frac{\sum_{i=1}^{n} \mathcal{X}_i}{n} - x| > \delta) + \varepsilon \\
\leq &2||f||_\infty \frac{x(1-x)}{n\delta^2} + \varepsilon \\
\leq &||f||_\infty \frac{1}{2n\delta^2} + \varepsilon \rightrightarrows 0
\end{aligned}$$

□