Means, Variances, and the Law of Large Numbers

Andrew Nobel

October 2024

Orientation

Setting: Observations $X_1, \ldots, X_n$ drawn independently from a distribution P

Standard terminology

1. Population. Refers to the distribution P and related quantities (e.g. mean and variance)

2. Sample. Refers to the observations $X_1, \ldots, X_n$ and related quantities (e.g. sample mean and sample variance)

Statistical Inference: What can the sample tell us about the population?

Sample Mean and Variance

Sample Mean

Recall: Expectation is a measure of center, variance is a measure of spread

Fact: Let $X_1, \ldots, X_n$ be iid and let $\overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$ be their sample mean. Then

1. $\mathbb{E}(\overline{X}_n) = \mathbb{E}X$
2. $\mathrm{Var}(\overline{X}_n) = \mathrm{Var}(X)/n$

Statistical Interpretation

- $\overline{X}_n$ is an unbiased estimate of $\mathbb{E}X$
- $\mathrm{Var}(\overline{X}_n) = \mathbb{E}(\overline{X}_n - \mathbb{E}X)^2$ is the mean squared error of $\overline{X}_n$
- Estimates $\overline{X}_n$ improve (less spread) as sample size n increases

Note: Fact and interpretation readily extend to sequences $f(X_1), \ldots, f(X_n)$

Examples

EX 1: If $X_1, \ldots, X_n$ are iid U(a, b) then

- $\overline{X}_n$ is an unbiased estimate of $\mathbb{E}X = (a + b)/2$
- MSE of $\overline{X}_n$ is $(b - a)^2/12n$

EX 2: Let $X_1, \ldots, X_n$ be iid. Fix $A \subseteq \mathbb{R}$ and define $U_i = \mathbb{I}(X_i \in A)$

- $U_1, \ldots, U_n$ are iid Bern(q) where $q = \mathbb{P}(X \in A)$
- $\overline{U}_n \stackrel{\mathsf{d}}{=} n^{-1}\text{Bin}(n, q)$
- $\overline{U}_n$ is an unbiased estimate of $\mathbb{E}U = q$
- MSE of $\overline{U}_n$ is $\text{Var}(U)/n = q(1 - q)/n$

Sample Variance

Definition: The sample variance of n observations $X_1, \ldots, X_n$ is given by

$$S_n^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \overline{X}_n)^2$$

Fact: If the observations are independent $\mathbb{E}S_n^2 = \mathrm{Var}(X)$. Thus S_n^2 is an unbiased estimate of variance

Note: Fact relies on the numerical identities

- $(\sum_{i=1}^n x_i)^2 = \sum_{i=1}^n \sum_{j=1}^n x_i x_j$
- $\sum_{i=1}^n (x_i - \overline{x}_n)^2 = \sum_{i=1}^n x_i^2 - n(\overline{x}_n)^2$

Markov and Chebyshev Inequalities

Markov's and Chebyshev's Inequalities

Markov's inequality: If $X \geq 0$ and $t > 0$ then

$$\mathbb{P}(X \geq t) \leq \frac{\mathbb{E}X}{t}$$

Chebyshev's Inequality: For each $t > 0$

$$\mathbb{P}(|X - \mathbb{E}X| \geq t) \leq \frac{\mathrm{Var}(X)}{t^2}$$

- Chebyshev bounds the probability that X is far from its expectation
- Upper bound is less than 1 if $t > \mathrm{SD}(X)$

Weak Law of Large Numbers

Theorem: Let $X_1, X_2, \ldots$ be iid with $\mathrm{Var}(X)$ finite. For each $t > 0$, as the sample size n tends to infinity

$$\mathbb{P}\left(\left| \frac{1}{n} \sum_{i=1}^{n} X_i \; - \; \mathbb{E}X \right| \geq t \right) \to 0$$

Terminology: The theorem says that the average of $X_1, \ldots, X_n$ *converges in probability* to the expected value of X as the sample size increases

Notation: We write $(1/n) \sum_{i=1}^{n} X_i \xrightarrow{p} \mathbb{E}X$

Weak Law of Large Numbers, Examples

Note: Can apply the theorem to any sequence $f(X_1), f(X_2), \ldots$ where $f : \mathbb{R} \to \mathbb{R}$

- $(1/n)\sum_{i=1}^n X_i^2 \xrightarrow{p} \mathbb{E}X^2$
- $(1/n)\sum_{i=1}^n \log(X_i) \xrightarrow{p} \mathbb{E}\log(X)$ if $X > 0$
- $(1/n)\sum_{i=1}^n \mathbb{I}(X_i \leq t) \xrightarrow{p} \mathbb{P}(X \leq t) = F_X(t)$

Cor: If $X_1, X_2, \ldots$ are iid then sample variance $S_n^2 \xrightarrow{p} \mathrm{Var}(X)$

Cauchy-Schwartz Inequality

Cauchy-Schwarz Inequality

Fact: If X and Y are random variables then

$$|\mathbb{E}(XY)| \leq \mathbb{E}|XY| \leq \sqrt{\mathbb{E}X^2}\sqrt{\mathbb{E}Y^2}$$

Immediate Corollaries

1. $|\mathrm{Cov}(X,Y)| \leq \sqrt{\mathrm{Var}(X)}\sqrt{\mathrm{Var}(Y)}$
2. $-1 \leq \mathrm{Corr}(X,Y) \leq 1$

Note: Inequality not restricted to expectations. If $g, h : \mathbb{R} \to \mathbb{R}$ then

$$\left|\int g(x)h(x)dx\right| \leq \sqrt{\int g(x)^2dx}\sqrt{\int h(x)^2dx}$$