9: Asymptotics

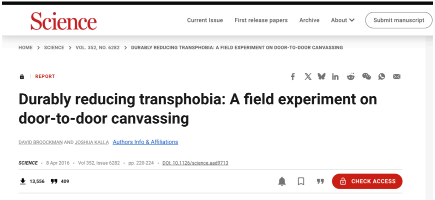
Naijia Liu

Spring 2025

Where are we and where to go?

- Last time: introducing estimators, looking at finite-sample properties.
- Now: can we say more as sample size grows?

Political canvassing study


Science

Current Issue First release papers Archive About Submit manuscript

HOME > SCIENCE > VOL. 352, NO. 6282 > DURABLY REDUCING TRANSPHOBIA: A FIELD EXPERIMENT ON DOOR-TO-DOOR CANVASSING

REPORT

Durably reducing transphobia: A field experiment on door-to-door canvassing

DAVID BROOCKMAN AND JOSHUA KALLA Authors Info & Affiliations

SCIENCE • 8 Apr 2016 • Vol 352, Issue 6282 • pp. 220-224 • DOI: 10.1126/science.aad9713

13,556 409 CHECK ACCESS

- Can canvassers change minds about topics like transgender rights?
- Experimental setting:
  - Randomly assign canvassers to have a conversation about transgender rights or a conversation about recycling.
  - Trans rights conversations focused on "perspective taking"
- Outcome of interest: support for trans rights policies.

Translating into math

- Outcome: $Y_i \in \{1 \text{ (least supportive)}, 2, 3, 4, 5 \text{ (most supportive)}\}$
- Treatment: $D_i \in \{0 \text{ (recycling script)}, 1 \text{ (trans rights script)}\}$
- Goal is to learn **something** about the joint distribution of (Y_i, D_i).
- Typical estimand would be the difference in conditional expectations:

$$\tau = \mathbb{E}[Y_i \mid D_i = 1] - \mathbb{E}[Y_i \mid D_i = 0]$$

- Typical plug in estimator would be the difference in sample means:

$$\hat{\tau}_n = \frac{\sum_{i=1}^n Y_i D_i}{\sum_{i=1}^n D_i} - \frac{\sum_{i=1}^n Y_i(1 - D_i)}{\sum_{i=1}^n (1 - D_i)}$$

- Today: what happens to the distribution of $\hat{\tau}_n$ as n grows?

Current knowledge

- For i.i.d. r.v.s, $X_1, \ldots, X_n$, with $\mathbb{E}[X_i] = \mu$ and $\mathbb{V}[X_i] = \sigma^2$ we know that:
  - $\bar{X}_n$ is **unbiased**, $\mathbb{E}[\bar{X}_n] = \mathbb{E}[X_i] = \mu$
  - Sampling variance is $\mathbb{V}[\bar{X}_n] = \frac{\sigma^2}{n}$ where $\sigma^2 = \mathbb{V}[X_i]$
  - None of these rely on a **specific distribution** for X_i!
- Assuming $X_i \sim \mathcal{N}(\mu, \sigma^2)$, we know the exact distribution of $\bar{X}_n$.
  - What if the data isn't normal? What is the sampling distribution of $\bar{X}_n$?
- **Asymptotics**: approximate the sampling distribution of $\bar{X}_n$ as n gets big.

Sequence of sample means

- What can we say about the sample mean n gets large?
- Need to think about sequences of sample means with increasing n:

$$\begin{aligned}
\overline{X}_1 &= X_1 \\
\overline{X}_2 &= (1/2) \cdot (X_1 + X_2) \\
\overline{X}_3 &= (1/3) \cdot (X_1 + X_2 + X_3) \\
\overline{X}_4 &= (1/4) \cdot (X_1 + X_2 + X_3 + X_4) \\
\overline{X}_5 &= (1/5) \cdot (X_1 + X_2 + X_3 + X_4 + X_5) \\
&\vdots \\
\overline{X}_n &= (1/n) \cdot (X_1 + X_2 + X_3 + X_4 + X_5 + \cdots + X_n)
\end{aligned}$$

- Note: this is a sequence of random variables!

Asymptotics and Limits

- Asymptotic analysis is about making **approximations** to finite sample properties.
- Useful to know some properties of deterministic sequences:

Definition

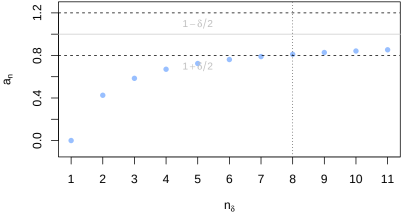
A sequence $\{a_n : n = 1, 2, \ldots\}$ has the **limit** a written $a_n \to a$ as $n \to \infty$ if for all $\delta > 0$ there is some $n_\delta < \infty$ such that for all $n \geq n_\delta$, $|a_n - a| \leq \delta$.

- a_n gets closer and closer to a as n gets larger (a_n **converges** to a)
- $\{a_n : n = 1, 2, \ldots\}$ is **bounded** if there is $b < \infty$ such that $|a_n| < b$ for all n.

Limit example: $(n{-}1)/n$

Definition

A sequence $\{a_n : n = 1, 2, \ldots\}$ has the **limit** a written $a_n \to a$ as $n \to \infty$ if for all $\delta > 0$ there is some $n_\delta < \infty$ such that for all $n \geq n_\delta$, $|a_n - a| \leq \delta$.



Convergence in Probability

Definition

A sequence of random variables, $\{Z_n : n = 1, 2, \dots\}$, is said to **converge in probability** to a value b if for every $\varepsilon > 0$,

$$\mathbb{P}(|Z_n - b| > \varepsilon) \to 0,$$

as $n \to \infty$. We write this $Z_n \xrightarrow{p} b$.

- Basically: probability that Z_n lies outside any (teeny, tiny) interval around b approaches 0 as $n \to \infty$.
- Economists write $\text{plim}(Z_n) = b$ if $Z_n \xrightarrow{p} b$.
- An estimator is **consistent** if $\hat{\theta}_n \xrightarrow{p} \theta$.
  - Distribution of $\hat{\theta}_n$ collapses on θ as $n \to \infty$.
  - Inconsistent estimators are bad bad bad: more data gives worse answers!

Law of large numbers

Weak Law of Large Numbers

Let $X_1, \ldots, X_n$ be a an i.i.d. draws from a distribution with mean $\mathbb{E}[|X_i|] < \infty$.
Let $\bar{X}_n = \frac{1}{n}\sum_{i=1}^n X_i$. Then, $\bar{X}_n \xrightarrow{p} \mathbb{E}[X_i]$.

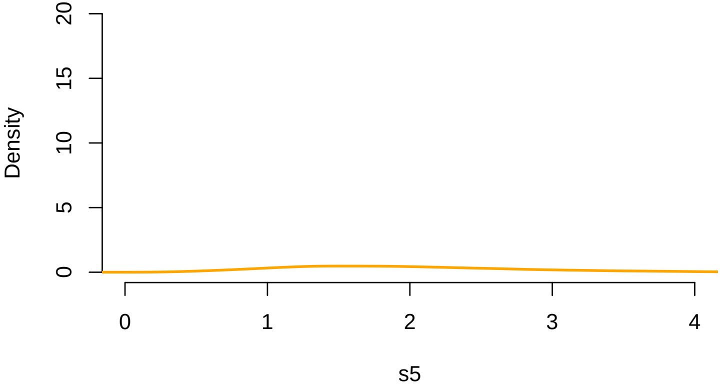
- Note: we don't assume finite variance, only finite expectation.
- Intuition: The probability of $\bar{X}_n$ being "far away" from μ goes to 0 as n gets big.
- Implies general consistency of **plug-in estimators**
  - If $\mathbb{E}[|g(X_i)|] < \infty$, then $\frac{1}{n}\sum_{i=1}^n g(X_i) \xrightarrow{p} \mathbb{E}[g(X_i)]$

LLN by simulation in R

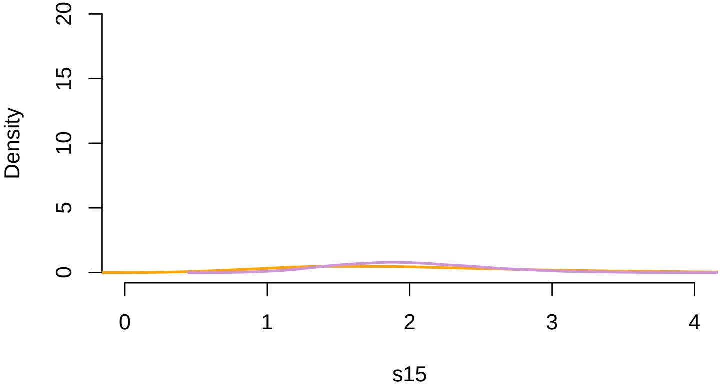
- Draw different sample sizes from Exponential distribution with rate 0.5
- $\mathbb{E}[X_i] = 2$

```r
set.seed(02140)
nsims <- 10000
holder <- matrix(NA, nrow = nsims, ncol = 6)
names_holder <- c("s5", "s15", "s30", "s100", "s1000", "s10000")
for (i in 1:nsims) {
  s5 <- rexp(n = 5, rate = 0.5)
  s15 <- rexp(n = 15, rate = 0.5)
  s30 <- rexp(n = 30, rate = 0.5)
  s100 <- rexp(n = 100, rate = 0.5)
  s1000 <- rexp(n = 1000, rate = 0.5)
  s10000 <- rexp(n = 10000, rate = 0.5)
  holder[i,1] <- mean(s5)
  holder[i,2] <- mean(s15)
  holder[i,3] <- mean(s30)
  holder[i,4] <- mean(s100)
  holder[i,5] <- mean(s1000)
  holder[i,6] <- mean(s10000)
}
```

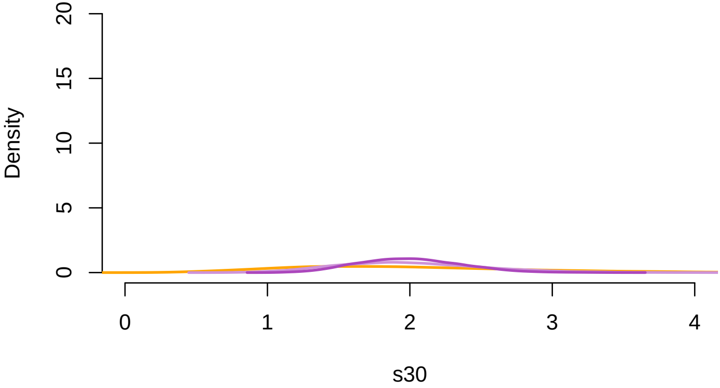
LLN in Simulation



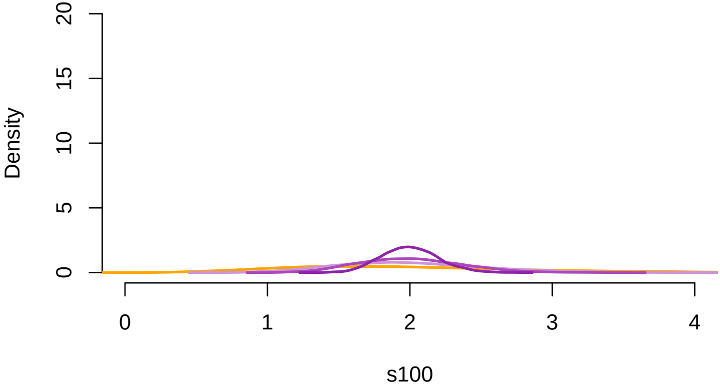
LLN in Simulation



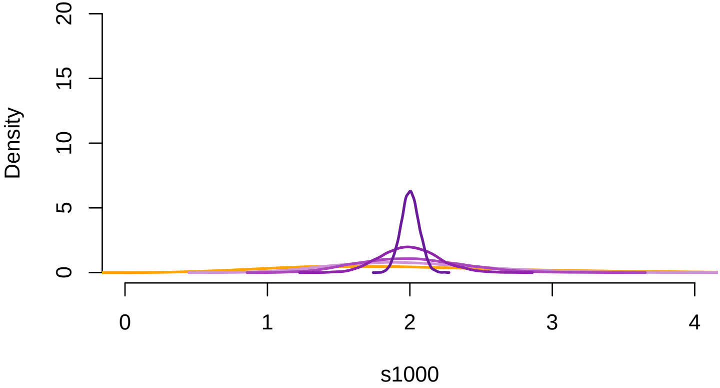
LLN in Simulation



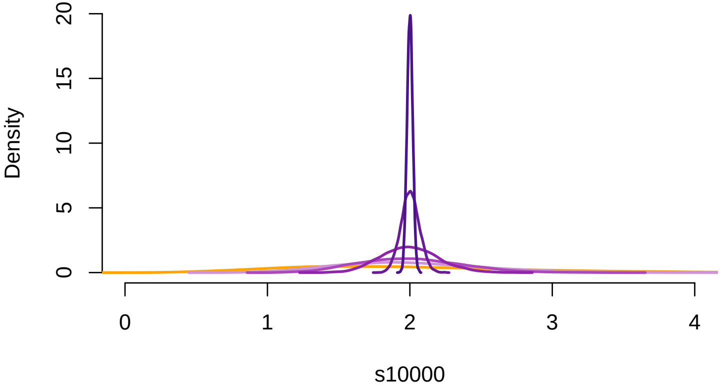
LLN in Simulation



LLN in Simulation



LLN in Simulation



Chebyshev Inequality

- How can we show convergence in probability? Can verify if we know specific distribution of $\hat{\theta}$.
- But can we say anything for arbitrary distributions?

Chebyshev Inequality

Suppose that X is r.v. for which $\mathbb{V}[X] < \infty$. Then, for every real number $\delta > 0$,

$$\mathbb{P}(|X - \mathbb{E}[X]| \geq \delta) \leq \frac{\mathbb{V}[X]}{\delta^2}.$$

- Variance places limits on how far an observation can be from its mean.

Proof of Chebyshev

- Let $Z = X - \mathbb{E}[X]$ with density $f_Z(x)$. Probability is just integral over the region:

$$\mathbb{P}(|Z| \geq \delta) = \int_{|x| \geq \delta} f_Z(x)\, dx$$

- Note that where $|x| \geq \delta$, we have $1 \leq x^2/\delta^2$, so

$$\mathbb{P}(|Z| \geq \delta) \leq \int_{|x| \geq \delta} \frac{x^2}{\delta^2} f_Z(x)\, dx \leq \int_{-\infty}^{\infty} \frac{x^2}{\delta^2} f_Z(x)\, dx = \frac{\mathbb{E}[Z^2]}{\delta^2} = \frac{\mathbb{V}[X]}{\delta^2}$$

- Under finite variance, applying this to $|\overline{X}_n - \mu|$ proves the LLN.

Properties of convergence in probability

1. **Continuous mapping theorem**: if $X_n \xrightarrow{p} c$, then $g(X_n) \xrightarrow{p} g(c)$ for any continuous function g.
2. if $X_n \xrightarrow{p} a$ and $Z_n \xrightarrow{p} b$, then
   - $X_n + Z_n \xrightarrow{p} a + b$
   - $X_n Z_n \xrightarrow{p} ab$
   - $X_n / Z_n \xrightarrow{p} a/b$ if $b > 0$

Thus, by LLN and CMT:

- $\left(\overline{X}_n\right)^2 \xrightarrow{p} \mu^2$
- $\log(\overline{X}_n) \xrightarrow{p} \log(\mu)$

Difference in means example

$$\hat{\tau}_n = \frac{\sum_{i=1}^n Y_i D_i}{\sum_{i=1}^n D_i} - \frac{\sum_{i=1}^n Y_i (1 - D_i)}{\sum_{i=1}^n (1 - D_i)}$$

- What about our difference in means estimator for the transphobia example?
- Let's take the sample mean for the treated units:

$$\frac{\sum_{i=1}^n Y_i D_i}{\sum_{i=1}^n D_i} = \frac{\frac{1}{n}\sum_{i=1}^n Y_i D_i}{\frac{1}{n}\sum_{i=1}^n D_i} \xrightarrow{p} \frac{\mathbb{E}[Y_i D_i]}{\mathbb{E}[D_i]} = \mathbb{E}[Y_i \mid D_i = 1]$$

  Last step uses iterated expectations and the fundamental bridge.

- Same idea for the other sample mean implies,

$$\hat{\tau}_n \xrightarrow{p} \mathbb{E}[Y_i \mid D_i = 1] - \mathbb{E}[Y_i \mid D_i = 0] = \tau$$

- Interpretation: Under iid sampling, adding more units gets us closer and closer to the truth.

Unbiased versus consistent

- By Chebyshev, unbiased estimators are consistent if $\mathbb{V}[\hat{\theta}_n] \to 0$.
- **Unbiased, not consistent**: "first observation" estimator, $\hat{\theta}^f_n = X_1$.
  - Unbiased because $\mathbb{E}[\hat{\theta}^f_n] = \mathbb{E}[X_1] = \mu$
  - Not consistent: $\hat{\theta}^f_n$ is constant in n so its distribution never collapses.
  - Said differently: the variance of $\hat{\theta}^f_n$ never shrinks.
- **Consistent, but biased**: sample mean with n replaced by $n-1$:

$$\frac{1}{n-1}\sum_{i=1}^{n} X_i = \frac{n}{n-1}\overline{X}_n \xrightarrow{p} 1 \times \mu$$

Consistent because $n/(n-1) \to 1$ as $n \to \infty$.

Multivariate LLN

- Let $\mathbf{X}_i = (X_{i1}, \ldots, X_{ik})$ be a random vectors of length k.
- Random (iid) sample of n of these k vectors, $\mathbf{X}_1, \ldots, \mathbf{X}_n$.
- Vector sample mean:

$$\overline{\mathbf{X}}_n = \frac{1}{n}\sum_{i=1}^{n} \mathbf{x}_i = \begin{pmatrix} \overline{X}_{n,1} \\ \overline{X}_{n,2} \\ \vdots \\ \overline{X}_{n,k} \end{pmatrix}$$

- **Vector WLLN**: if $\mathbb{E}[\|\mathbf{X}\|] < \infty$, then as $n \to \infty$, $\overline{\mathbf{X}}_n \xrightarrow{p} \mathbb{E}[\mathbf{X}]$.
  - Converge in probability of a vector is just convergence of each element.
  - $\mathbb{E}[\|\mathbf{X}\|] < \infty$ is equivalent to $\mathbb{E}[|X_{ij}|] < \infty$ for each $j = 1, \ldots, k$