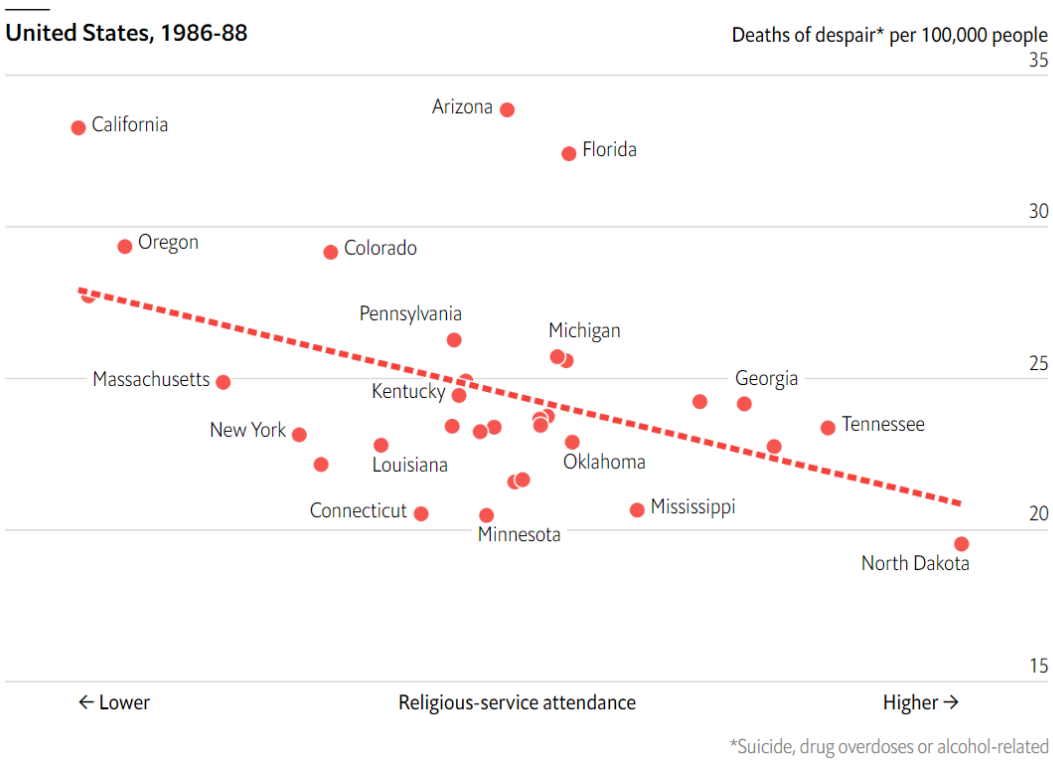
PUBH 526 Randomized Trials

Homework 1

1. The following graph was published in The Economist on February 27, 2023:



Source: "Places with high religious participation have fewer deaths of despair" https://www.economist.com/graphic-detail/2023/02/27/places-with-high-religious-participation-have-fewer-deaths-of-despair

Note: While the article is provided with this homework assignment, you have full access to this and other publications through Purdue Libaries. Try finding the article yourself, if you don't already know how.

a. Explain as if you're speaking to a general audience what message this graph is trying to convey. (1 point)

b. Can we conclude that increasing religious service attendance would reduce mortality due to suicide, drug overdoses, or excessive alcohol use? Why or why not? (1 point)

2. Explain all the assumptions made by this statistical model (4 points):

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$$
$$\varepsilon_i \sim \text{iid } N(0,\sigma^2)$$

3. 2300 participants were enrolled in a randomized trial. They were then randomly and equally allocated to one of two study arms. Over the course of the study, each arm lost approximately 45% of participants to attrition (loss to follow-up). Explain specifically how this may be a threat to both the internal and external validity of the trial. Hint: consider why loss to follow-up may occur: is it at random? If not, are the reasons the same or different between the two arms? (6 points)

Homework 1

Alexandra Chang

PUBH 526: Design and Analysis of Randomized Trials in Public Health

September 4, 2025

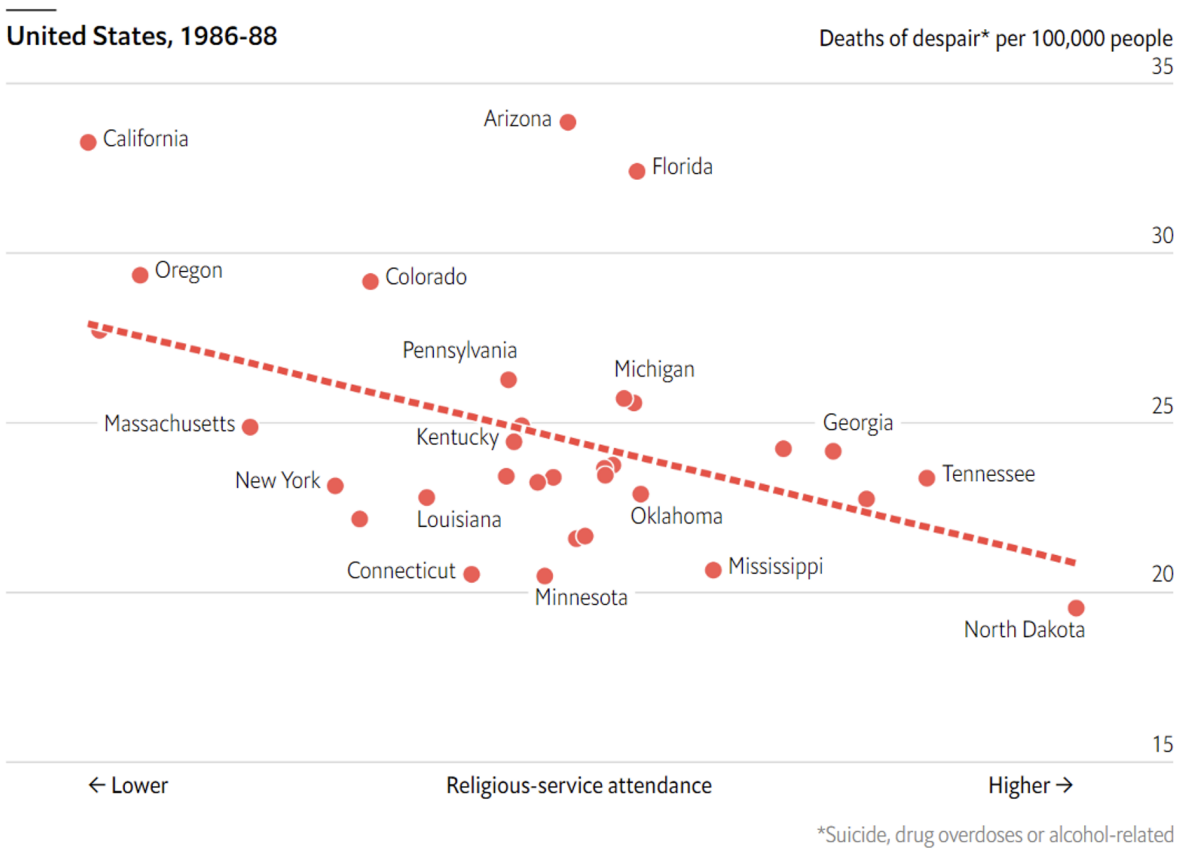


Figure 1: **Deaths of despair were highest in states with thinly attended churches**

*Suicide, drug overdoses or alcohol-related. †Data not available for every year. Source: "Opiates of the masses? Deaths of despair and the decline of American religion," Giles, Hungerman, and Oostrom, NBER Working Paper, 2023.

Question 1 (1 point)

Explain as if you're speaking to a general audience what message this graph is trying to convey.

- This graph indicates that U.S. states with more people attending **religious services tend to have fewer deaths** caused by suicide, drug overdoses, and alcohol use, which is what researchers refer to as "**deaths of despair**." The message is that being part of a religious community might protect people from dying from these causes, possibly because of the social support such communities provide.

Question 2 (1 point)

Can we conclude that increasing religious service attendance would reduce mortality due to suicide, drug overdoses, or excessive alcohol use? Why or why not?

- **No**, we cannot conclude this with certainty because the study is **observational**, not a randomized experiment. Although the studies show a **correlation** between higher church attendance and fewer deaths of despair, they do not prove that one causes the other. Other factors, such as community strength, local culture, economic conditions, or the legalization of alcohol sales, may also influence both religious attendance and health outcomes.

Question 3 (4 points)

Explain all the assumptions made by this statistical model:

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i, \quad \varepsilon_i \sim \text{iid } N(0, \sigma^2)$$

1. **Linearity:** The relationship between the predictor X_i (e.g., religious attendance) and the outcome Y_i (e.g., deaths of despair) is assumed to be linear.

2. **Independence:** Each observation (i.e., each state or person) is independent of the others. One data point doesn't influence another.

3. **Homoscedasticity:** The variability (or variance) of the errors ε_i is the same across all values of X_i. This means the spread of the data around the regression line stays consistent.

4. **Normality of Errors:** The errors ε_i are normally distributed with a mean of 0. This is important for making statistical inferences like confidence intervals and hypothesis tests.

5. **Identically Distributed Errors (i.i.d.):** All error terms come from the same distribution, reinforcing that their behavior is consistent across observations.

Question 4 (6 points)

2300 participants were enrolled in a randomized trial. They were then randomly and equally allocated to one of two study arms. Over the course of the study, each arm lost approximately 45% of participants to attrition (loss to follow-up). Explain specifically how this may be a threat to both the internal and external validity of the trial. Hint: consider why loss to follow-up may occur: is it at random? If not, are the reasons the same or different between the two arms?

Internal Validity Threats

- If the **reasons for attrition** differ between arms, the remaining participants might no longer be comparable. For example, if sicker people drop out more in one arm, any differences in outcomes could be due to this imbalance, not the treatment.
- Even if attrition is balanced in number, it could be differential in cause, which leads to **selection bias**. This then undermines the ability to attribute outcomes solely to the intervention.
- If attrition were truly random, validity would be less threatened. However, attrition is often **non-random** (e.g., related to treatment side effects or participant burden), which would directly compromise causal inference.

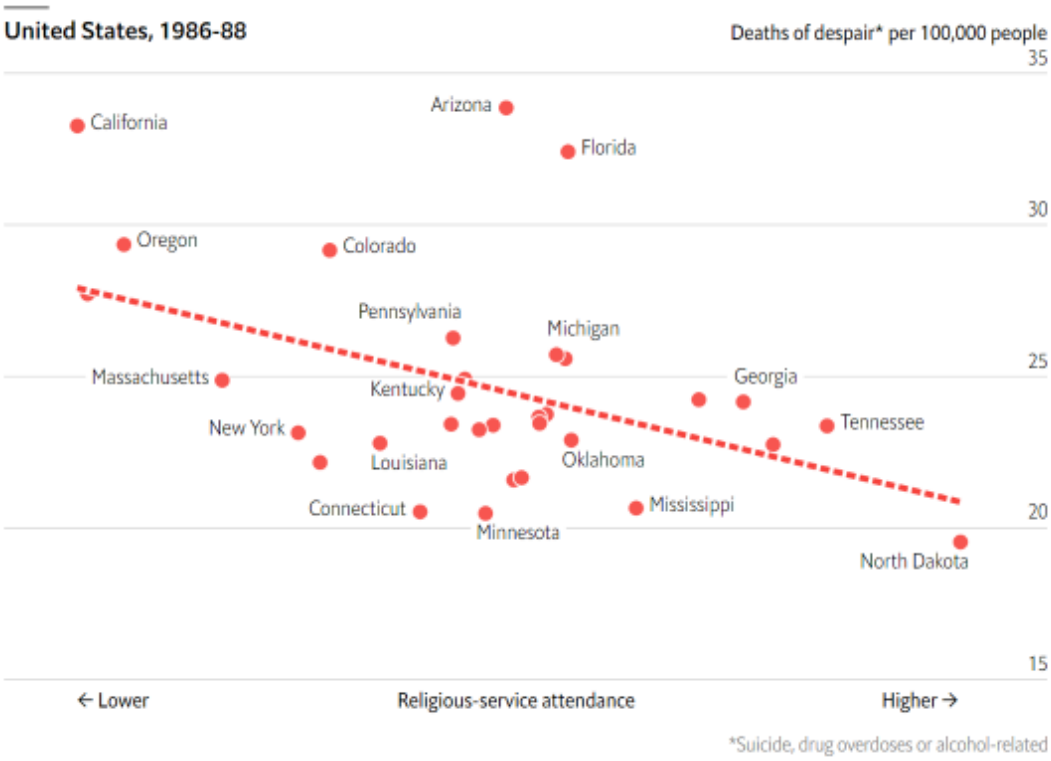
External Validity Threats

- Losing 45% of participants means the study results may no longer represent the **original target population**.
- If those who stayed in the study differ meaningfully from those who dropped out (e.g., healthier, more motivated), the results may not **generalize** to real-world settings or broader groups.
- High attrition reduces the confidence in applying results to policy or practice, since the intervention's **feasibility** and **acceptability** in broader populations is unclear when nearly half of participants cannot be retained.

PUBH 526 Randomized Trials

Homework 1

1. The following graph was published in The Economist on February 27, 2023:



Source: "Places with high religious participation have fewer deaths of despair" https://www.economist.com/graphic-detail/2023/02/27/places-with-high-religious-participation-have-fewer-deaths-of-despair

Note: While the article is provided with this homework assignment, you have full access to this and other publications through Purdue Libaries. Try finding the article yourself, if you don't already know how.

a. Explain as if you're speaking to a general audience what message this graph is trying to convey. (1 point)

   US states with higher attendance of religious services had lower rates of death from or related to suicide, drug overdoses, or alcohol.

b. Can we conclude that increasing religious service attendance would reduce mortality due to suicide, drug overdoses, or excessive alcohol use? Why or why not? (1 point)

No – we can't assume from this regression alone that there is a causal relationship between religious service attendance and mortality rate specifically due to "deaths of despair". Moreover, we can't conclude that this relationship, observed at a state level, would also hold for individual people (i.e., we can't conclude that if a person increases their religious service attendance, they are less likely to die from a "death of despair") – this is an example of an ecological fallacy.

2. Explain all the assumptions made by this statistical model (4 points):

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$$
$$\varepsilon_i \sim \text{iid } N(0,\sigma^2)$$

The observations and their errors are independent.
There is a linear relationship between X and Y.
All errors are normally distributed with a constant variance, σ^2.

3. 2300 participants were enrolled in a randomized trial. They were then randomly and equally allocated to one of two study arms. Over the course of the study, each arm lost approximately 45% of participants to attrition (loss to follow-up). Explain specifically how this may be a threat to both the internal and external validity of the trial. Hint: consider why loss to follow-up may occur: is it at random? If not, are the reasons the same or different between the two arms? (6 points)

If the loss to follow-up is completely at random, internal and external validity are preserved, although statistical power is reduced due to the smaller sample size.

If the loss to follow-up is systematic but similar across the two arms (e.g., people drop out for the same reasons in each arm), internal validity may still be preserved because the comparison between the arms remains reasonable – the only systematic difference between them is which treatment is applied, so it can be reasonable to attribute any differences in outcomes to the differences between treatments. However, external validity may be compromised, as the participants who remain in the study may not be fully representative of the original or target population.

If the loss to follow-up is systematic and differs between study arms, then both internal and external validity are threatened. The two arms are different in more ways than the treatments applied to them, so it can now be difficult to attribute any observed differences in outcomes between the two arms specifically to the treatments being investigated (i.e., internal validity is reduced). Similarly, as people systematically drop out of your study, your sample may no longer resemble or represent the full target population, reducing external validity.