Simulated Overlap for d = 0.08

Anthony Steven Dick

Contents

Load Required Libraries

```
library(dplyr)
library(tidyr)
library(ggplot2)
library(ggthemes)
```

Simulate Data for Cohen's d = 0.08

```
effect_size <- 0.08

# Gryffindor color palette
gryffindor_red <- "#7F0909"
gryffindor_gold <- "#D3A625"
gryffindor_red_dark <- "#6A0D0D"
gryffindor_red_light <- "#A61D1D"

# Define color scale
scale_color_rpsy <- scale_color_manual(values = c(
  "Unexposed" = gryffindor_red,
  "PCE" = gryffindor_gold,
  "Unexposed_overlap" = gryffindor_red_dark,
  "PCE_overlap" = gryffindor_red_light
))

# Simulate distributions with Monte Carlo integration
set.seed(4443451)
overlap_fun <- function(x) pmin(dnorm(x, 0, 1), dnorm(x, effect_size, 1))

n <- 10000
d <- data.frame(x = runif(n, -3, 3.5),
                PCE = runif(n, 0, 0.4),
                Unexposed = runif(n, 0, 0.4)) %>%
  mutate(
    PCE = ifelse(PCE <= dnorm(x, 0, 1), PCE, NA),
```

```
    Unexposed = ifelse(Unexposed <= dnorm(x, effect_size, 1), Unexposed, NA)
  )

d_long <- d %>%
  pivot_longer(cols = -x, names_to = "Distribution", values_to = "y") %>%
  mutate(
    overlap = ifelse(y <= overlap_fun(x), paste0(Distribution, "_overlap"), Distribution),
    overlap = factor(overlap)
  ) %>%
  filter(!is.na(y))
```

Figure 6A: Two Distributions by Group

```
d_long %>%
  ggplot(aes(x, y, color = Distribution)) +
  geom_point(alpha = 0.5, size = 1.3) +
  facet_wrap(~ Distribution, ncol = 1) +
  labs(title = "Distribution of Outcomes in the Non-Exposed and PCE Groups",
       subtitle = "Cohen's d = 0.08; Each point represents one person") +
  scale_color_rpsy +
  theme_bw() +
  theme(
    text = element_text(size = 18),
    panel.border = element_blank(),
    panel.grid.minor = element_blank(),
    axis.line.x = element_line(),
    axis.line.y = element_line(),
    axis.title = element_blank()
  )
```

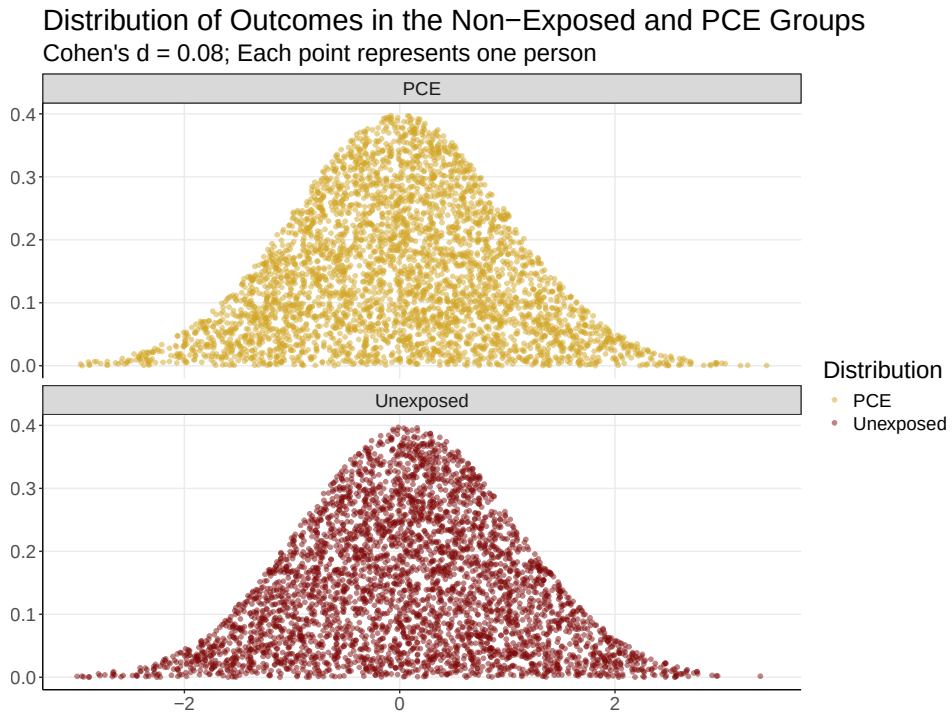

Figure 6B: Overlap Visualization

```
plot <- d_long %>%
  ggplot(aes(x, y, color = Distribution)) +
  geom_point(alpha = 0.5, size = 1.3) +
  labs(title = "Overlap of Two Distributions", subtitle = "Cohen's d = 0.08") +
  scale_color_rpsy +
  theme_bw() +
  theme(
    text = element_text(size = 18),
    panel.border = element_blank(),
    panel.grid.minor = element_blank(),
    axis.line.x = element_line(),
    axis.line.y = element_line(),
    axis.title = element_blank()
  )

plot
```

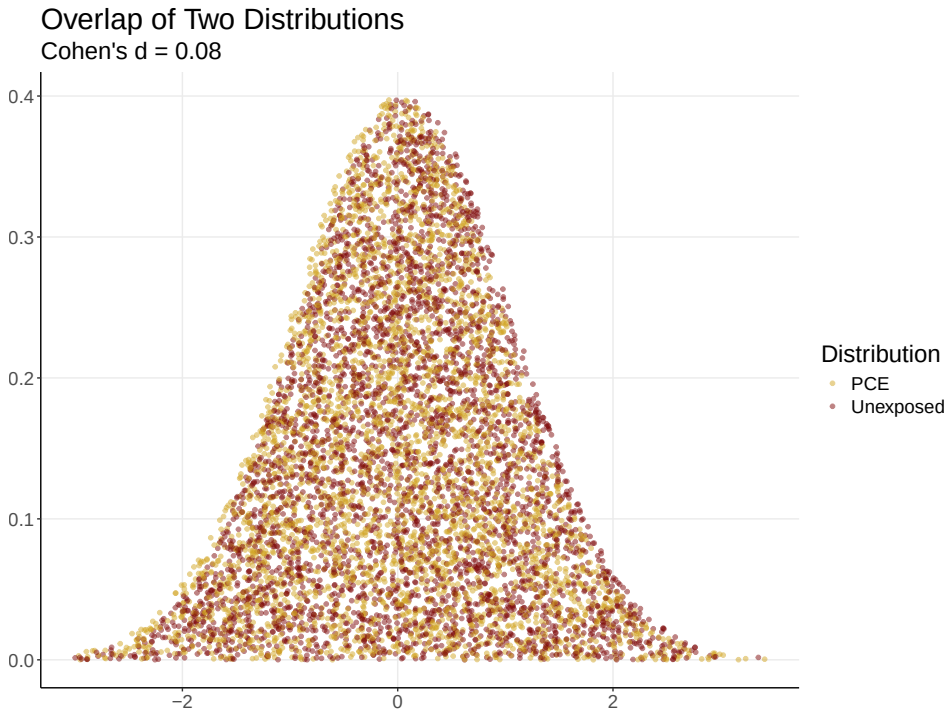

Figure 6C: Annotated Percent Overlap

```
labels <- d_long %>%
  group_by(overlap) %>%
  summarise(n = n(), .groups = "drop") %>%
  mutate(
    prop = paste0(round(n / sum(n) * 100, 1), "%"),
    x = c(1.5, 0.25, -1, 0.25),
    y = c(0.2, 0.25, 0.2, 0.2)
  )

d_long %>%
```

```r
ggplot(aes(x, y, color = overlap)) +
geom_point(alpha = 0.5) +
geom_label(data = labels, aes(x = x, y = y, label = prop, color = overlap),
           vjust = "center", show.legend = FALSE, size = 5) +
labs(title = "Percentage of Observations in Each Area",
     subtitle = "Frequency Understanding of Overlap") +
scale_color_rpsy +
theme_bw() +
theme(
  text = element_text(size = 18),
  panel.border = element_blank(),
  panel.grid.minor = element_blank(),
  axis.line.x = element_line(),
  axis.line.y = element_line(),
  axis.title = element_blank()
)
```