The interval that spans two IPC phases

GAM is 14.9%, at the top of IPC Phase 3. The design-adjusted interval runs 11.1% to 18.7%, which spans Phase 3 and Phase 4 — and the survey cannot say which the population is in.

Pierre Robentz Cassion

Lesson 6 of 8

Nutrition Analysis: CMAM and SMART — Judging a survey

What this lesson covers

- The number the survey exists to produce
- The estimate, with the design
- The IPC phases
- What that means, said plainly
- What would have happened without the design effect
- Read it with the plausibility report
- The full reporting block
- What comes next

The number the survey exists to produce

- Everything so far has been preparation.

The estimate, with the design — In Python (cont.)

```
import numpy as np
import pandas as pd

analysable = smart[smart["whz"].between(-5, 5)].copy()
analysable["gam"] = (analysable["whz"] < -2) | (analysable["oedema"] == True)
analysable["sam"] = (analysable["whz"] < -3) | (analysable["oedema"] == True)


def cluster_prevalence(df, column, cluster="cluster"):
    n = len(df)
    p = df[column].mean()

    totals = df.groupby(cluster)[column].agg(["sum", "size"])
    residual = totals["sum"] - p * totals["size"]
    m = len(totals)
```

The estimate, with the design — In Python (cont.)

```
    variance = m / (m - 1) * (residual**2).sum() / n**2
    se = np.sqrt(variance)
    srs = p * (1 - p) / n
    return {"p": p, "se": se, "deff": variance / srs, "clusters": m, "n": n}


for column in ["gam", "sam"]:
    r = cluster_prevalence(analysable, column)
    t = 2.045                                  # t(0.975, 29 df)
    print(f"{column.upper()}: {r['p']:.1%}  95% CI {r['p'] - t * r['se']:.1%} to "
          f"{r['p'] + t * r['se']:.1%}  deff {r['deff']:.2f}  n {r['n']}")
```

The estimate, with the design — In R

```
library(survey)

design <- svydesign(ids = ~cluster, weights = NULL, data = analysable)

svyciprop(~I(whz < -2 | oedema), design, method = "logit")
svymean(~I(whz < -2 | oedema), design, deff = TRUE)
```

The estimate, with the design

	Estimate	95% CI	Design effect	n
GAM	14.9%	**11.1 – 18.7%**	2.29	852
SAM	3.8%	2.3 – 5.2%	1.20	852

The estimate, with the design

- The design effect is 2.29 for GAM and 1.20 for SAM — Two indicators, one survey, two design effects — because...
- The naive interval would be 12.5 – 17.3% — Ignoring the clustering makes the interval 37% narrower, and the next...

The IPC phases

Phase	GAM by weight-for-height
1 — Acceptable	under 5%
2 — Alert	5 to 9.9%
3 — Serious	10 to 14.9%
4 — Critical	**15 to 29.9%**
5 — Extremely critical	30% and above

The IPC phases — In Python

```
def ipc_phase(gam):
    for threshold, phase in [(0.05, 1), (0.10, 2), (0.15, 3), (0.30, 4)]:
        if gam < threshold:
            return phase
    return 5


r = cluster_prevalence(analysable, "gam")
t = 2.045
low, high = r["p"] - t * r["se"], r["p"] + t * r["se"]

print(f"point estimate: phase {ipc_phase(r['p'])}")
print(f"interval spans: phase {ipc_phase(low)} to phase {ipc_phase(high)}")
```

The IPC phases — In R

```
c(point = 0.149, low = 0.111, high = 0.187)
```

The IPC phases

- The point estimate is 14.9%, which is Phase 3 — by one tenth of a point
- The interval runs 11.1% to 18.7%, which spans Phase 3 and Phase 4

What that means, said plainly

Global acute malnutrition is estimated at 14.9% (95% CI 11.1–18.7). The point estimate falls in IPC Phase 3 (Serious). The confidence interval spans Phase 3 and Phase 4 (Critical), so **this survey cannot determine which phase the population is in.**

What that means, said plainly

- It refuses to round to the threshold — 14.9% reported as "about 15%" has made a phase classification by rounding, and...
- It states the interval before the phase — A reader who sees "Phase 3" first will not revise it when they reach the...
- It says the survey cannot decide — That is the three-branch verdict the survey course insisted on, and here the third...

What would have happened without the design effect — In Python

```python
srs_se = np.sqrt(r["p"] * (1 - r["p"]) / r["n"])
print(f"naive interval: {r['p'] - 1.96 * srs_se:.1%} to {r['p'] + 1.96 * srs_se:.1%}")
```

What would have happened without the design effect — In R

```
p <- 0.149; n <- 852
c(p - 1.96 * sqrt(p * (1 - p) / n), p + 1.96 * sqrt(p * (1 - p) / n))
```

What would have happened without the design effect

- **The clustering decides the classification whenever the estimate is within

Read it with the plausibility report

- **The inflated standard deviation (1.22) raises GAM.** A wider distribution puts more children past a fixed cut-off.
- **Team 3's low measurements raise GAM.** Its clusters give 22.3% against 9.2 to 16.2% for the others, and it...

Read it with the plausibility report — In Python

```python
without_team3 = analysable[analysable["team"] != 3]
r3 = cluster_prevalence(without_team3, "gam")
print(f"excluding team 3: {r3['p']:.1%} on n={r3['n']}, "
      f"{r3['clusters']} clusters")
```

Read it with the plausibility report — In R

```
analysable |> filter(team != 3) |> summarise(gam = mean(gam), n = n())
```

Read it with the plausibility report

- Report the sensitivity, do not substitute it — The headline stays the full-sample estimate; the exclusion goes beside...

The full reporting block — Example

```
Global acute malnutrition (weight-for-height z < -2 or oedema)
  14.9%   95% CI 11.1-18.7   n = 852   design effect 2.29   30 clusters
  IPC Phase 3 by point estimate; interval spans Phase 3 and Phase 4.

Severe acute malnutrition (z < -3 or oedema)
  3.8%    95% CI 2.3-5.2     n = 852   design effect 1.20

Plausibility: 4 of 6 SMART checks failed (SD 1.22, digit preference in one team,
age heaping 24%, team bias -0.42 to -1.09). Two of the four push the estimate
upward. Excluding the affected team gives 12.6% on 628 children.

Analysable denominator: 852 of 930 measured. 66 had no computable z-score
(32 missing weight or age, 30 outside the reference range, 4 heights in metres,
recoverable) and 12 were flagged by the WHO bounds.
```

What comes next

- The survey says what the situation is.

Read the full lesson, with runnable code

https://data-analysis.cassion.dev/courses/nutrition-analysis-cmam-smart/nutrition-06-prevalence-and-ipc