Sampling Variances

- Typically, θ (theta) is used to denote a **“true” effect size**
- We use θ_k to denote the true effect of some specific study k.
- But this true effect size is **not identical** with the **observed effect size** that we find in the published results!
- We use a “hat” symbol (^) to clarify that all we have is **only an estimate** of the true effect size.
- The observed effect in study k can therefore be written as $\hat{\theta}_k$.

Sampling Variances

True Effect Size (θ_k)

Estimand

Effect Size Formula

Ingredients

- 150g unsalted butter, plus extra for greasing
- 150g plain chocolate, broken into pieces
- 150g plain flour
- ½ tsp baking powder
- ½ tsp bicarbonate of soda
- 200g light muscovado sugar
- 2 large eggs

Method

1. Heat the oven to 160C/140C fan/gas 3. Grease and base line a 1 litre heatproof glass pudding basin and a 450g loaf tin with baking parchment.
2. Put the butter and chocolate into a saucepan and melt over a low heat, stirring. When the chocolate has all melted remove from the heat.

Estimator

Observed Effect Size ($\hat{\theta}_k$)

Estimate

→ Why does $\hat{\theta}_k$ differ from θ_k?

credits to Simon Grund

Sampling Variances

Why does $\hat{\theta}_k$ differ from θ_k?

→ Because of **sampling error** ε_k

- In every primary study, researchers can only draw **a small sample** from the **whole population**.
  - E.g., when we want to examine benefits of regular exercise on the cardiovascular health of primary care patients, we will only be able to include a small selection of patients, not *all* primary care patients in the world.
- The fact that a study can only take small samples from an infinitely large population means that the **observed effect will differ** from the **true population effect**.

Sampling Variances

In Notation:

$$\hat{\theta}_k = \theta_k + \varepsilon_k$$

- It is obviously desirable for $\hat{\theta}_k$ to be as close to the true effect size as possible, and for $\boldsymbol{\varepsilon_k}$ **to be minimal**.
- All things equal, studies with a smaller ε provide a **more precise** estimate of the true effect size
- Meta-analysis takes this into account: we give studies with a **greater precision** (less sampling error) a **greater weight**.

Sampling Variances

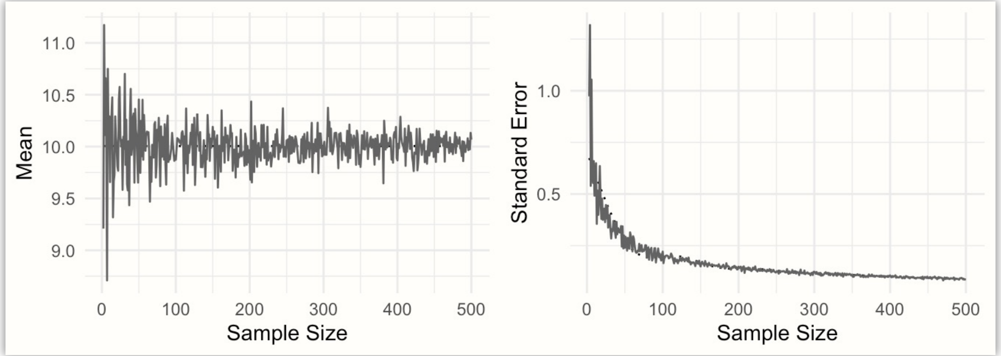
- Therefore, we need to know **how large the expected sampling error** of an observed effect size is.
- However, θ_k is unknown, so ε_k is also unknown.
- However, we can use statistical theory to **<u>approximate</u>** the sampling error
- A common way to **quantify $\boldsymbol{\varepsilon}$** is through the estimated **standard error** (SE), which is the square root of the sampling variance $\sqrt{V}=$ SE.
- The standard error is defined as the **standard deviation** of the **sampling distribution**.
- A sampling distribution is the **distribution of a metric** we get when we **draw random samples** with the same sample size n from our population **many, many times**.

Sampling Variances

- Using “closed-form” formulas to calculate the sampling variance, we could **quite accurately estimate the standard error using <u>only</u> the sample we have at hand**.
- The standard error of the mean depends on the sample size of a study.

→ As n becomes larger, the standard error becomes smaller, meaning that a study’s estimate of the true population mean becomes more precise.

$$SE_{\bar{X}} = \frac{\bar{X}}{\sqrt{n}}$$



Sampling Variances

These are the quintessential elements we need to conduct a meta-analysis. The:

1. **observed effect size**, and its
2. **precision**, expressed as the **standard error**.

→ If these two types of information can be calculated from a published study, it is usually also possible to perform a meta-analytic synthesis!

→ This applies not only to means, but to all commonly used effect size measures (e.g. standardized mean differences, correlations, odds ratios, …)