Lot Quality Assurance Sampling (LQAS)

hypothesis testing is more nuanced that involves calculating precise prevalence estimates and confidence intervals, testing whether a particular prevalence or something more extreme is probable or not

LQAS is a classification method that answers a yes/no question of whether a lot's prevalence is meeting a set threshold value or not

priority is getting the extremes really well-classified, okay with units between p_l and p_u not being correctly classified

α and β error levels are how often you're okay with making mistakes at either extremes

<u>Benefits</u>

answering a simple question so reduced sample size

simple classification rules that require little statistical training to implement

classifications are very actionable and can lead directly to program management decisions

often used in decentralized data collection and program management

helps understand local variation because done on a very small geographic scale

can be aggregate data from LQAS to get local and regional estimates and be part of complex sampling technique for national population prevalence

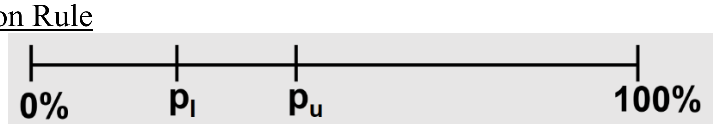
Two-Way Classification

Step 1	Randomly select n individuals in an area
Step 2	Count the number x that has trait of interest
Step 3	compare x to decision rule, d and classify a lot in low or high category

if $x < d$, lot is in low category

if $x \geq d$, lot is in high category

<u>Determining Decision Rule</u>



β-error = probability of classifying as high when it's actually low

α-error = probability of classifying as low when it's actually high

p_u = upper limit proportion
  less than α of being classified in the upper category
  at least $1 - \alpha$ chance to be classified in the lower category

p_l = lower limit proportion
  less than β of being classified in the upper category
  at least $1 - \beta$ chance to be classified in the lower category

search through a large range of sample sizes to find the n and d that satisfy the constraints using binomial distribution to calculate probabilities

for a given n, determine if there is a single value d that satisfies both p_u and p_l with given α and β error levels

once $n \geq 100$, LQAS becomes similar to complex sampling
increase accuracy of classification with smaller α and β error levels by increasing sample size
decrease distance between categories with tighter p_l and p_u range by increasing sample size

Example Decision Rule

$$n = 19, d = 9$$

If less than 9 out of 19 students per class are wearing pink, that class is classified in the low category. If 9 or more students per class are wearing pink, that class is classified in the high category.

$$p_u = 60\%, \alpha = 0.10$$

At 60% pink outfits, we want at least 90% probability of that class being classified in the upper category and less than 10% probability of that class being classified in the lower category.

$$p_l = 30\%, \beta = 0.10$$

At 30% pink outfits, we want at least 90% probability of that class being classified in the lower category and less than 10% probability of that class being classified in the upper category.