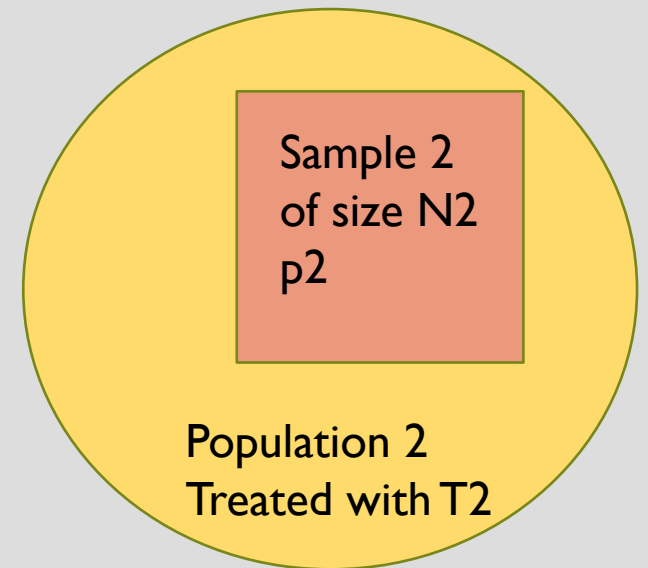
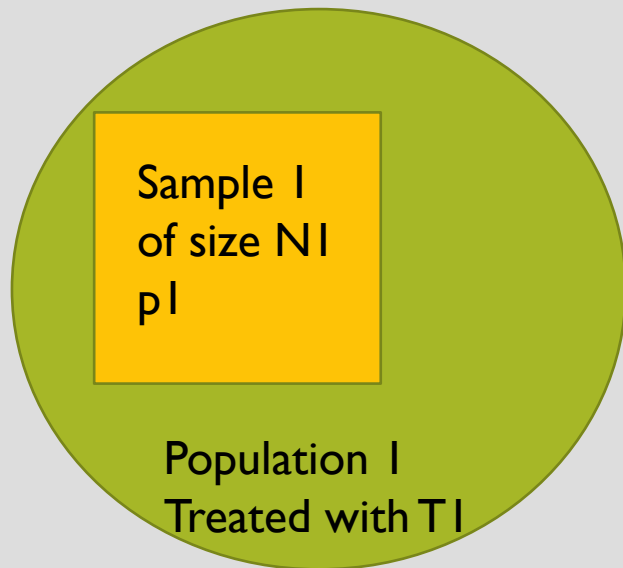


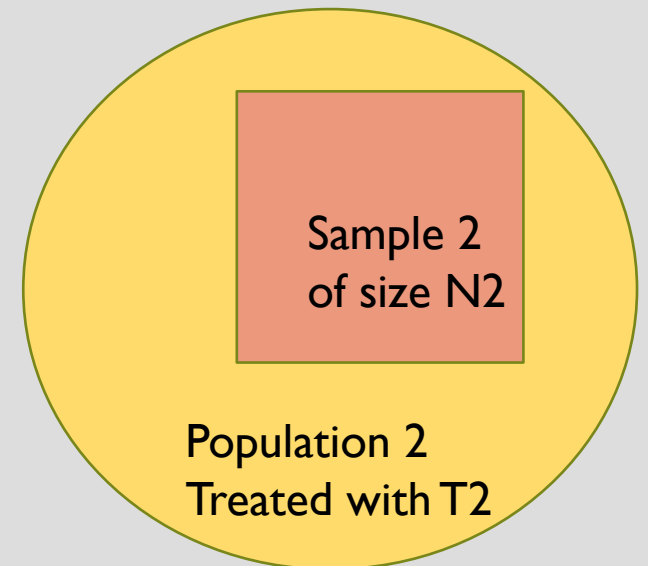
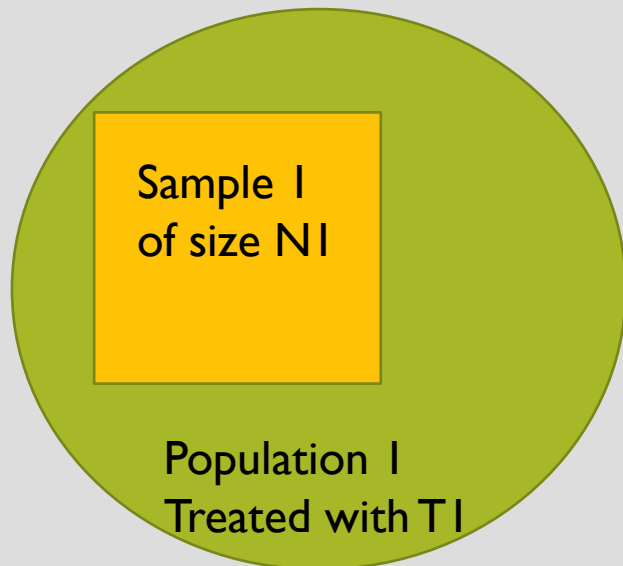
IS TREATMENT T1 BETTER THAN T2?

- Give T1 to N_1 patients and T2 to N_2 patients
 - Often $N_1 = N_2 = N$
- Record success rates p_1 of T1 and p_2 of T2



IS TREATMENT T1 BETTER THAN T2?

- Give T1 to N1 patients and T2 to N2 patients
 - Often $N1 = N2 = N$
- Record success rates $p1$ of T1 and $p2$ of T2
- Statistics view to the Question
 - 1. Is $p1$ **statistically significantly** larger than $p2$?
 - I.e. is “ $p1 > p2$ ” likely to hold also in other samples?
 - 2. How much **uncertainty** are there in the estimated success rates?
 - I.e. where estimates from other samples would be likely to reside?
- Crucial Q for study design:
 - 3. **How should we choose N1 and N2?**



HYPOTHESIS TESTING

- **H₀ (NULL HYPOTHESIS):** There is no difference in success rates between the treatments
- **H_A (ALTERNATIVE HYPOTHESIS):** There is a difference in success rates between the treatments
- **Significance level α :** “Reject H₀” and “accept H_A” if P-value (calculated assuming H₀) is $< \alpha$
 - If α is defined before the experiment, then the proportion of false rejection of H₀ would be α in repeated experiments
 - By making α small (say 0.00001) we can protect from false positive findings (Type I errors) but increase false negative findings (Type II errors)
 - By keeping α larger (say 0.05) we have more statistical power to reject H₀ but we are more likely to make a false positive finding (Type I error)

Suomeksi:

Merkitsevyytaso

Hylkää / hyväksy

Tyypin I / II virhe

HYPOTHESIS TESTING FRAMEWORK

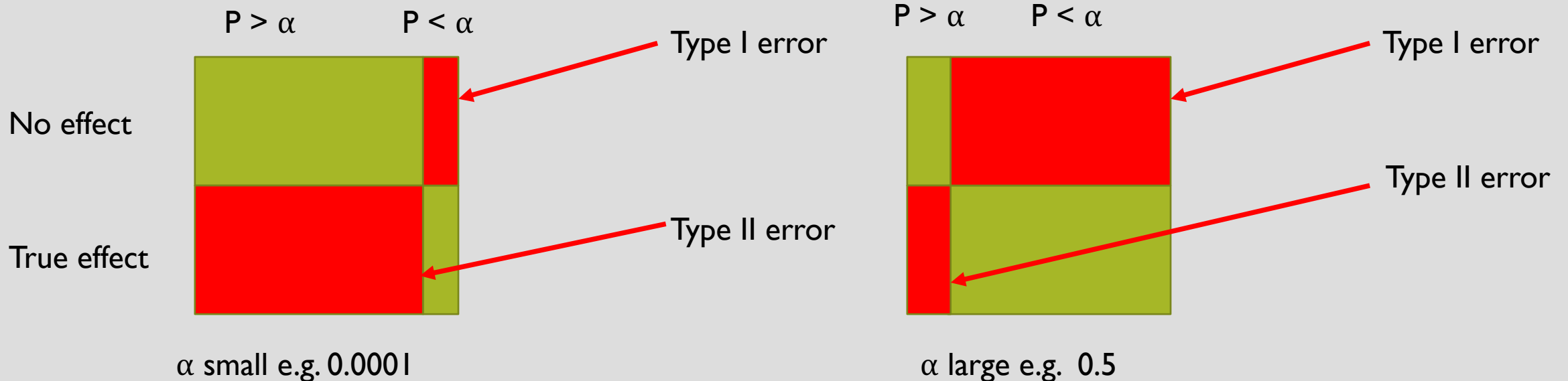
- Type I error: seeing an effect that is not real, a false positive
- Type II error: failing to see an effect that is real, a false negative
- Significance level α affects how likely different types of errors are to occur

RESULT FROM THE TEST

		RESULT FROM THE TEST	
		$P > \alpha$: H0	$P < \alpha$: HA
TRUE STATE OF THE WORLD	No effect		Type I error
	True effect	Type II error	

HYPOTHESIS TESTING FRAMEWORK

- Type I error: seeing an effect that is not real, a false positive
- Type II error: failing to see an effect that is real, a false negative
- With small significance level we rarely make Type I errors, but may often make Type II errors
- With large significance level we rarely make Type II errors, but may often make Type I errors



INTERPRET PROP.TEST() FROM HYPOTHESIS TESTING FRAMEWORK

```
# (ii)  
prop.test(A,N,p.null, correct = FALSE)
```

```
##  
## 1-sample proportions test without continuity correction  
##  
## data: A out of N, null probability p.null  
## X-squared = 17.195, df = 1, p-value = 3.373e-05  
## alternative hypothesis: true p is not equal to 0.081  
## 95 percent confidence interval:  
## 0.08910332 0.10551545  
## sample estimates:  
## p  
## 0.097
```

A = 485
N = 5000
p.null = 0.081

QUESTIONS

- Assume you are designing a study of the effect of a new treatment by a randomized clinical trial
- What is the null hypothesis?
- What is the alternative hypothesis?
- What is relationship between P-value and significance level?
- What does it mean if we do a Type I error?
- What does it mean if we do a Type II error?
- How would you choose the significance level?
 - That is, which properties of your study affect your choice and what consequences does your choice have?

STATISTICAL VS CLINICAL SIGNIFICANCE

confidence intervals
↔

Null value |



- No evidence for departure from null – do not change clinical practice



- May be an important effect, but sample size is too small to be sure – do a larger study



- Statistically significantly different from null value – but not clinically important



Range of clinical indifference ↔

- Statistically and clinically important – change practice