Prepare answers for Tuesday:

What is the **standard error**?

Be able to give a complete explanation to your neighbors

Hint 1: This is a TRUE value

What are the two formulas for the standard error?

Hint 2: For **Direct** and **Indirect** designs

Hint 3: It is the same for an estimate of a **mean** or a **treatment effect**

What is the formula for **estimating** the standard error from data?

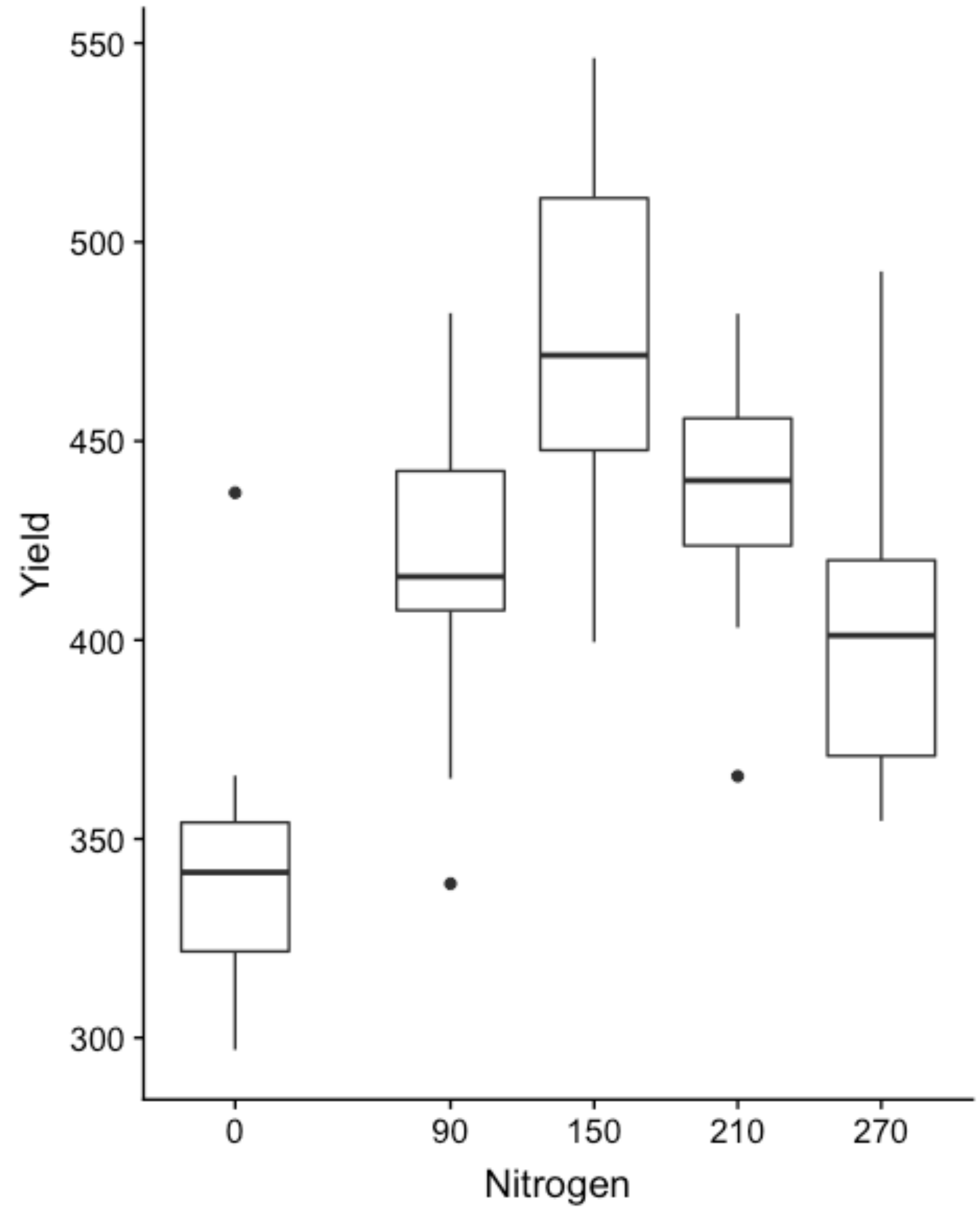
Hint 4: The calculation has 2 steps and uses sample size(s) twice

An experiment was run to evaluate effects of increased nitrogen fertilization on tuber yield of frying potatoes

5 nitrogen regimes (applied to plots): 0, 90, 150, 210, 270 lbs / acre at emergence

10 reps / treatment combination

Response: total yield per plot



What is the maximum improvement we could get?

Answer: What is the effect of the best level of Nitrogen?

Report $\hat{\delta}_{150-0} \pm t_c \times SED$

Can any addition of Nitrogen actually increase yield?

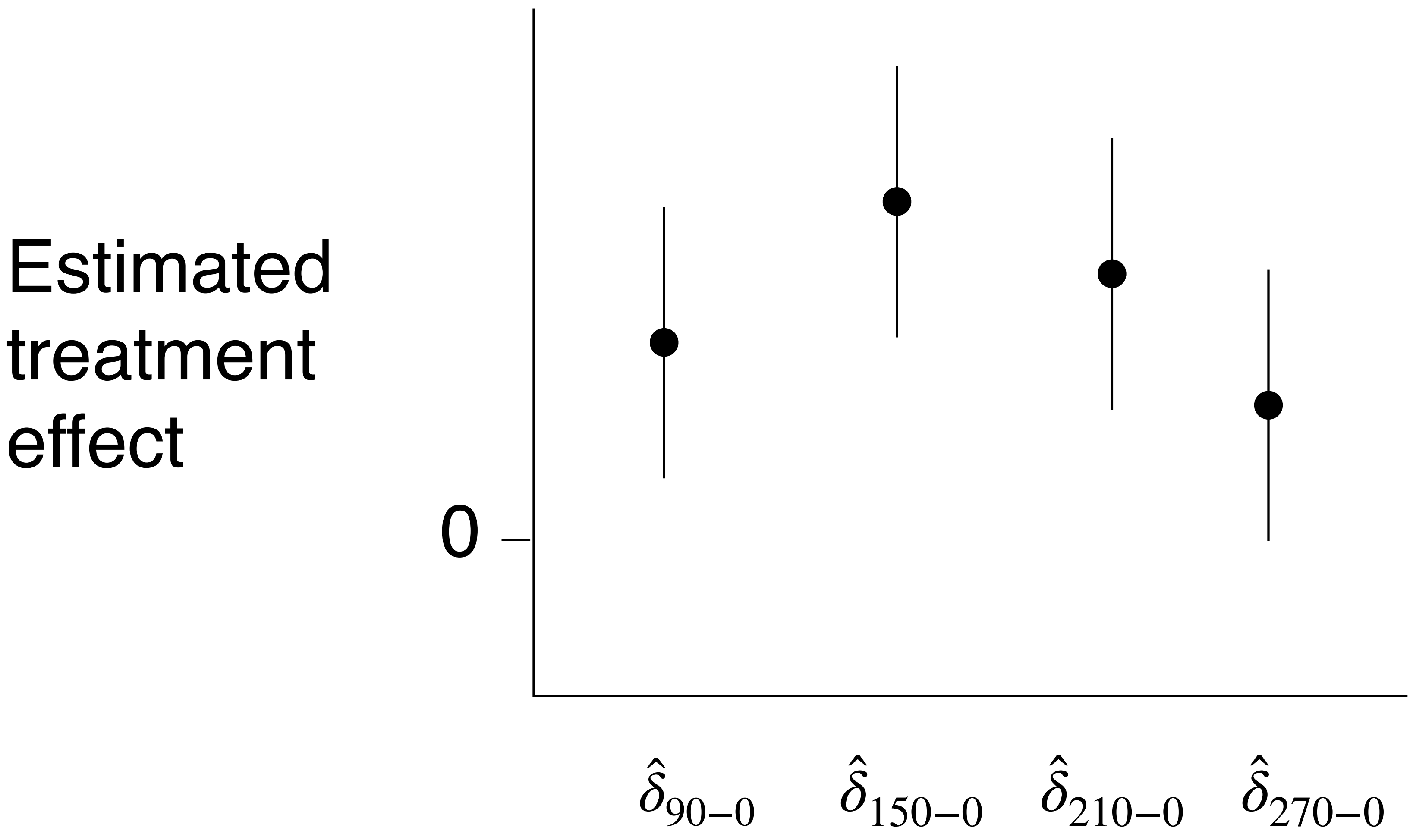
Answer: Power is highest when δ is biggest

Run T-test for $\delta_{150-0} = 0$

Both of these answers are misleading because we chose to run the statistics **because** the **estimated yield** for N-150 was highest

Once you look at the data, neither CIs nor p-values are valid

With 4 **new** treatments each compared to the control we are making 4 estimates



For each CI, there's a 5% chance that the TRUE effect is outside the interval

There is a 15% chance that the TRUE effect is outside **at least one** of these intervals

What is the maximum improvement we could get?

There’s a good chance the biggest estimated effect was over-estimated

We’re “safe” if we can ensure all CIs include their TRUE values

Strategy: Adjust CIs so that the chance that **any** true effect is outside of the interval is $100\alpha\,\%$

CI: $\hat{\delta}_{i-0} \pm t_c^D \times SED$

$t_{\alpha,df}^{D(k)}$ comes from the Dunnett distribution

k: # **new** treatments (excluding control)

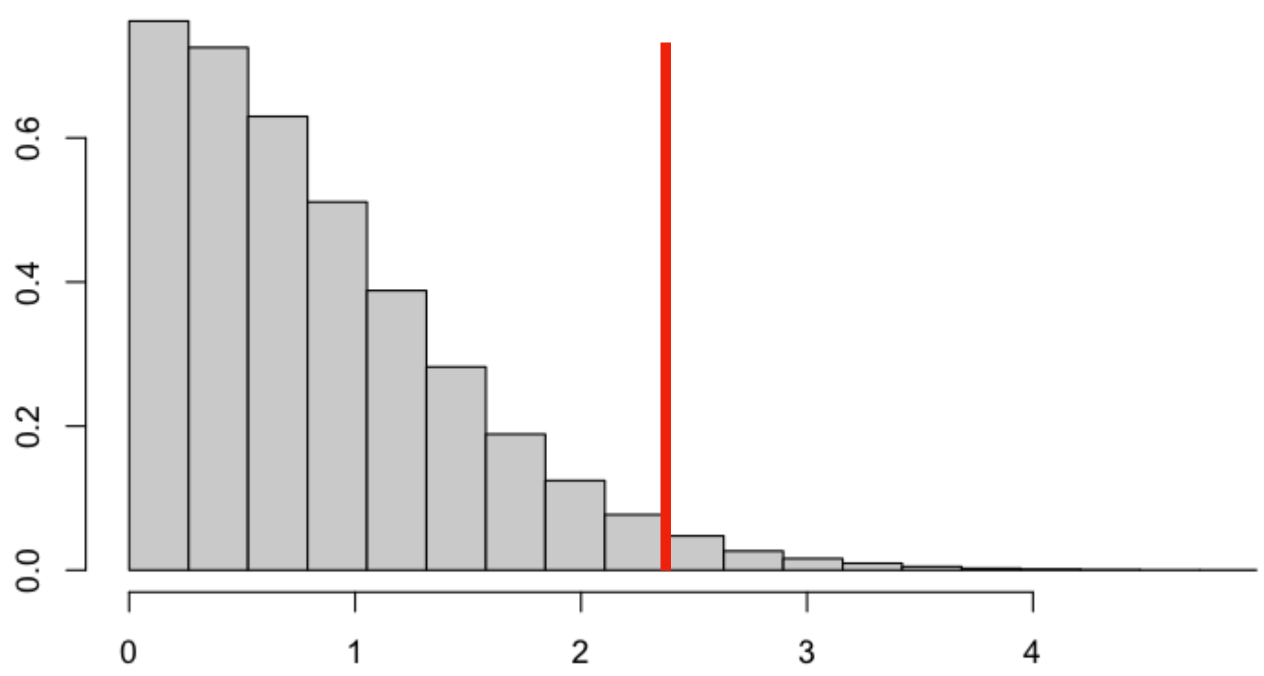
α: False Positive rate

df: Degrees of freedom from all treatments

Bigger value than t_c

Accounting for **multiple comparisons**

T-distribution



Estimate: $\hat{\delta}$, SED

$$\frac{|\hat{\delta} - \delta|}{\text{SED}}$$

Actual error

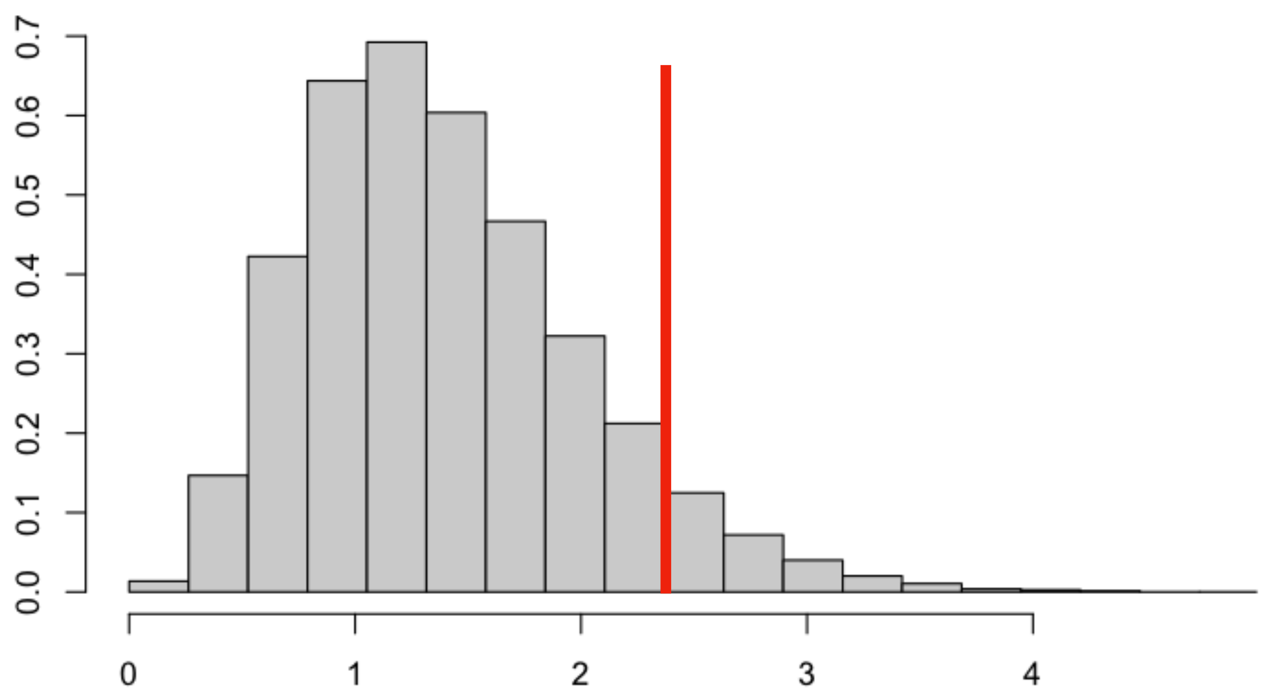
Estimated average error

Distribution of Normalized errors

How much bigger than SED could my actual error have been?

*Valid for a single treatment effect

Dunnett(4) distribution



Estimate: $\hat{\delta}_{90-0}$, $\hat{\delta}_{150-0}$, $\hat{\delta}_{210-0}$, $\hat{\delta}_{270-0}$, SED

$$\frac{max|\hat{\delta}_i - \delta_i|}{\text{SED}}$$

Actual size of **biggest** error

Estimated average error

Distribution of Biggest Normalized errors

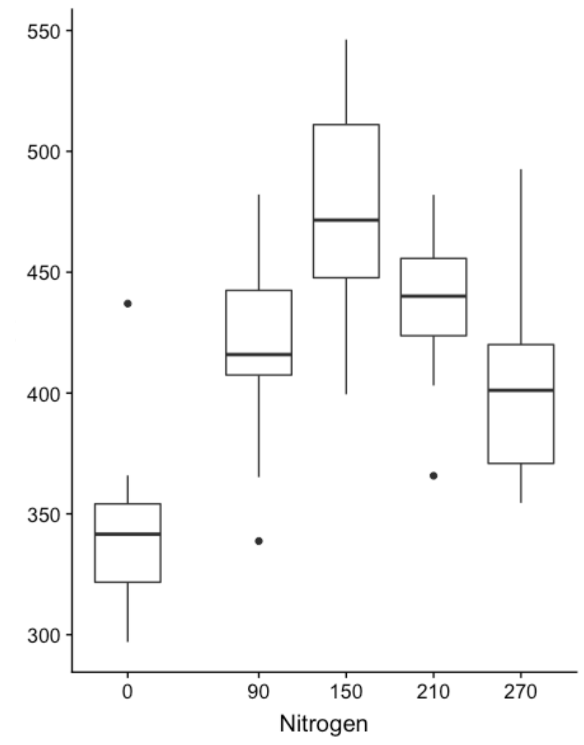
Can any addition of Nitrogen actually increase yield?

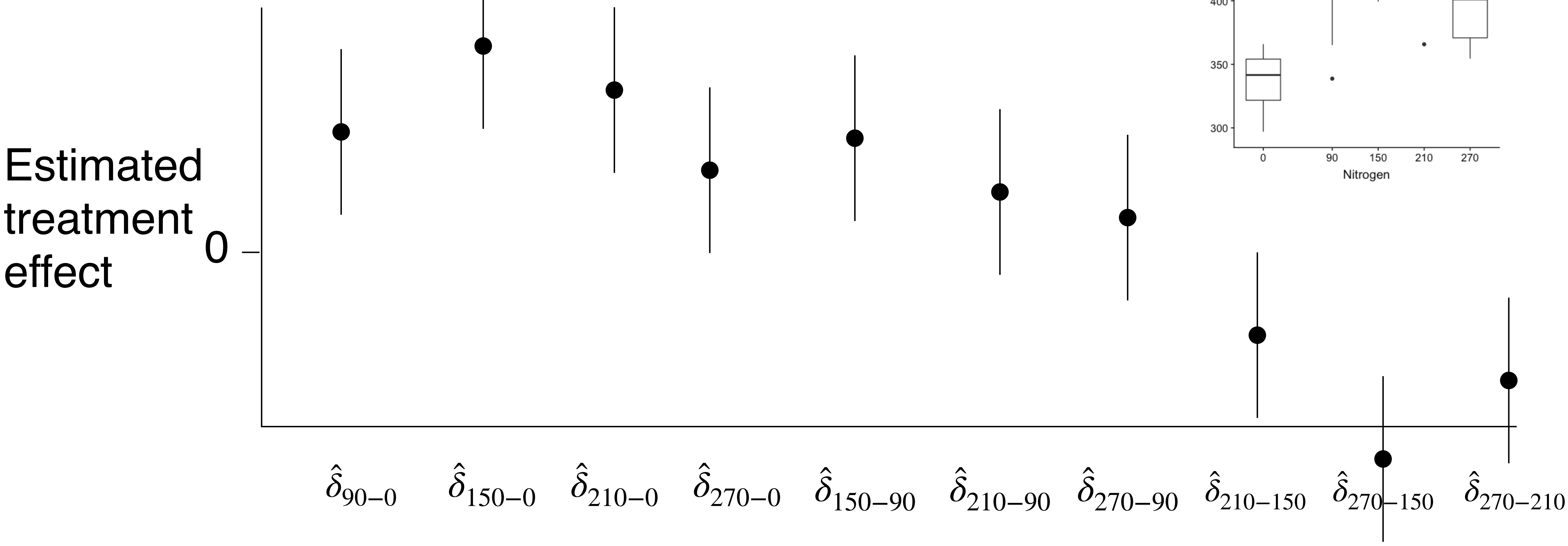
If we run a T-test for $\delta_{150-0} = 0$, the p-value will be **too small**

Instead calculate p-value from the Dunnett(4) distribution

Corrected p-value: Probability of **the biggest observed Normalized effect** being this large if all 4 treatments had no effect

Does **any** level of Nitrogen addition affect yield?





With 5 treatment levels, we can make 10 pairwise comparisons

At least 1 CI won't include the TRUE effect ~27% of the time

We'd conclude that at least 2 levels differ ~27% of the time even if Nitrogen had no effect at all

Solution:

CI: $\hat{\delta}_{i-0} \pm t_c^T \times SED$

$t_{\alpha,df}^{T(k)}$ comes from the Tukey distribution

k: # **total** treatments

α: False Positive rate

df: Degrees of freedom from all treatments

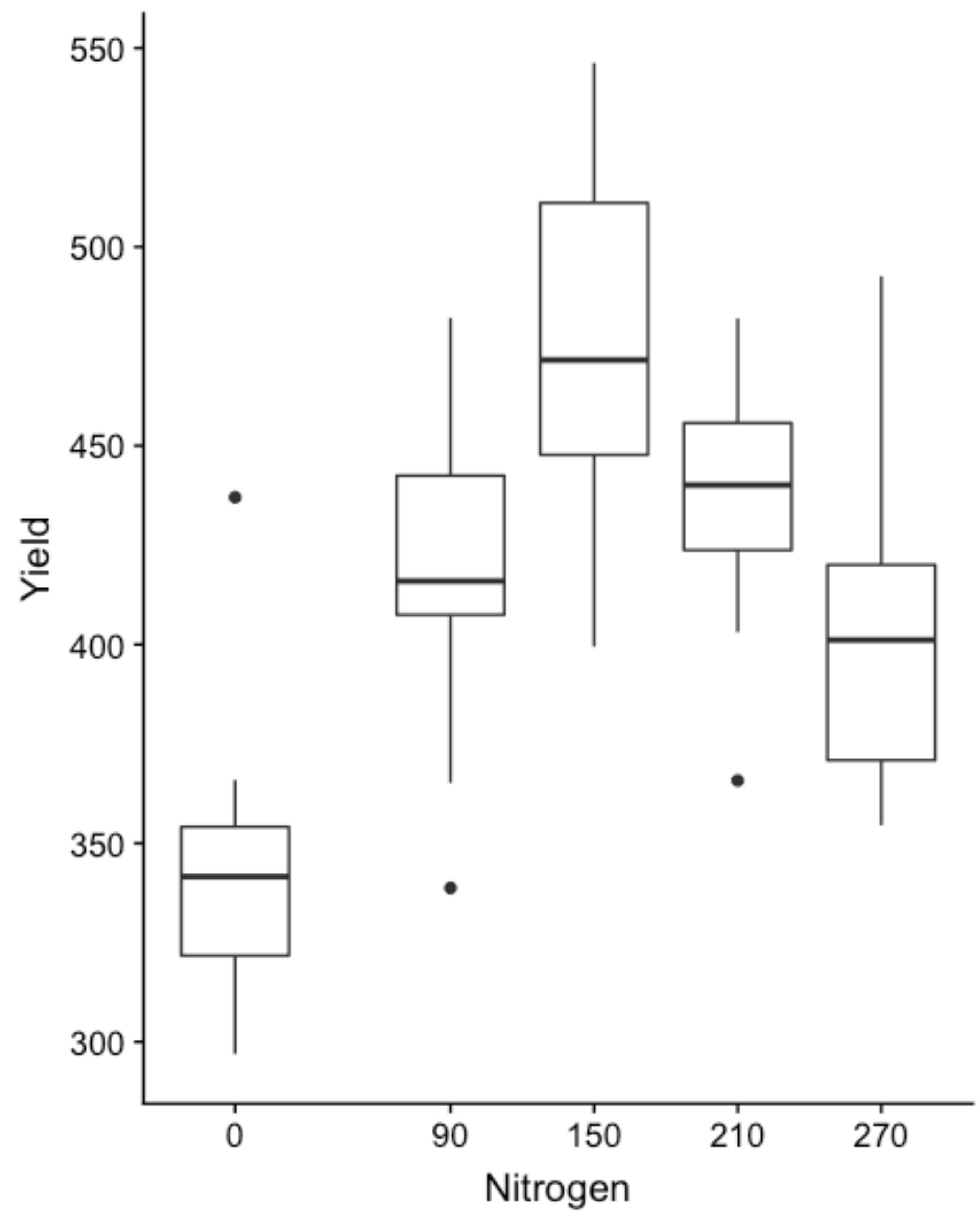
Bigger value than t_c from either T or Dunnett

An experiment was run to evaluate effects of increased nitrogen fertilization on tuber yield of frying potatoes

5 nitrogen regimes (applied to plots): 0, 90, 150, 210, 270 lbs / acre at emergence

10 reps / treatment combination

Response: total yield per plot



If you are specifically asked about the effect of N=90 vs N=0:

Use the T-distribution

Specify: emmeans(…, at=list(Nitrogen=c(0,90))

If you are interested in which (if any) are different from the control (N=0)

Use the Dunnett distribution

Cannot compare the new treatments

Remember: Not significant $\neq$ No effect

Specify: contrast(…,method = ‘trt.vs.ctrl’,ref=‘0’)

If you are interested in which (if any) are different from any other

Use the Tukey distribution

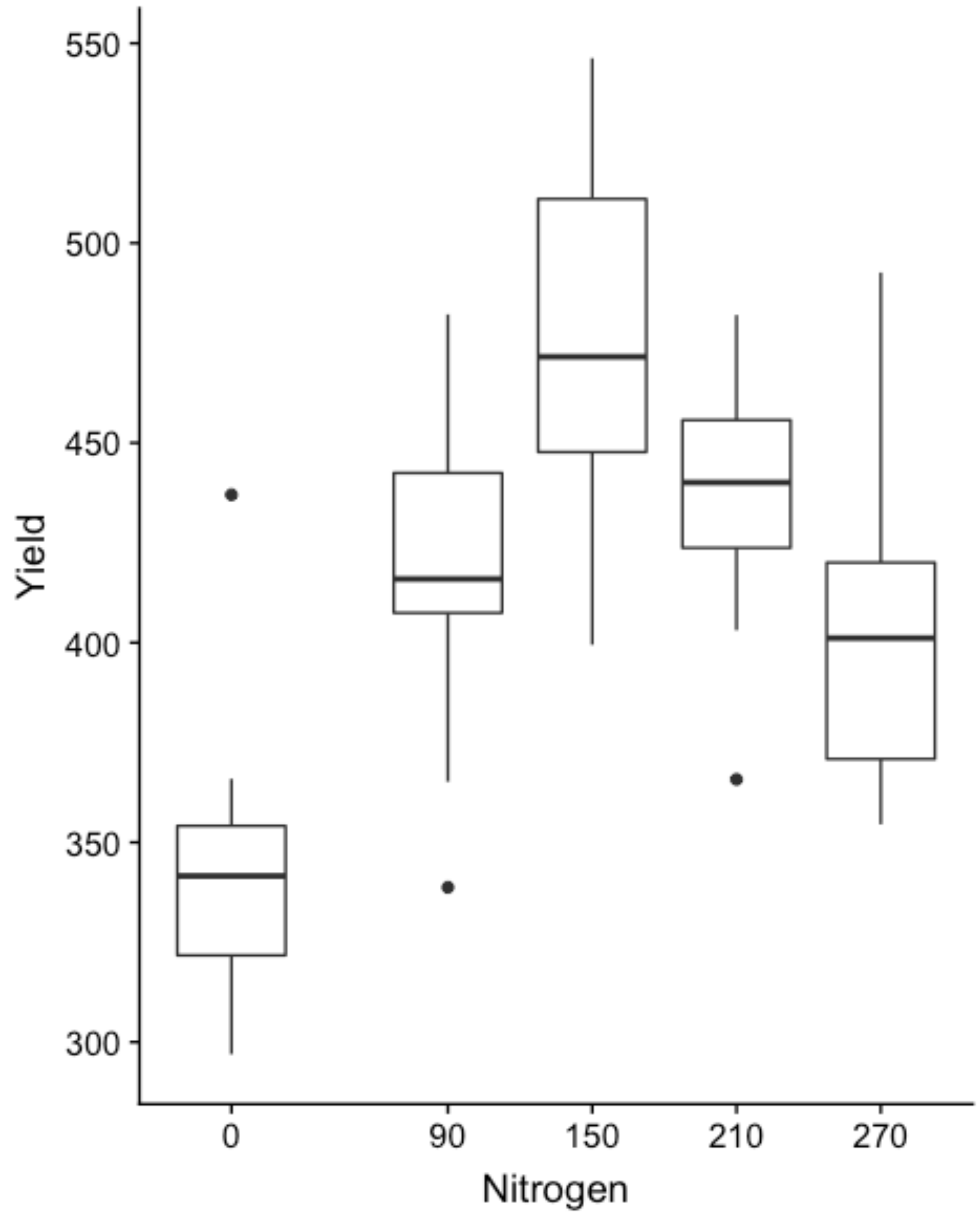
Most common situation

Specify: contrast(…,method = ‘pairwise’)

Analysis of Variance (ANOVA)

```
Analysis of Variance Table

Response: Yield
          Df Sum Sq Mean Sq F value    Pr(>F)    
Nitrogen   4  92042 23010.6  13.868 1.901e-07 ***
Residuals 45  74664  1659.2                      
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```



Describes the calculations used in a F-test

Addresses the question: Are there any treatment means that are not the same?

H_0: $\mu_1 = \mu_2 = \ldots = \mu_K$

Method:

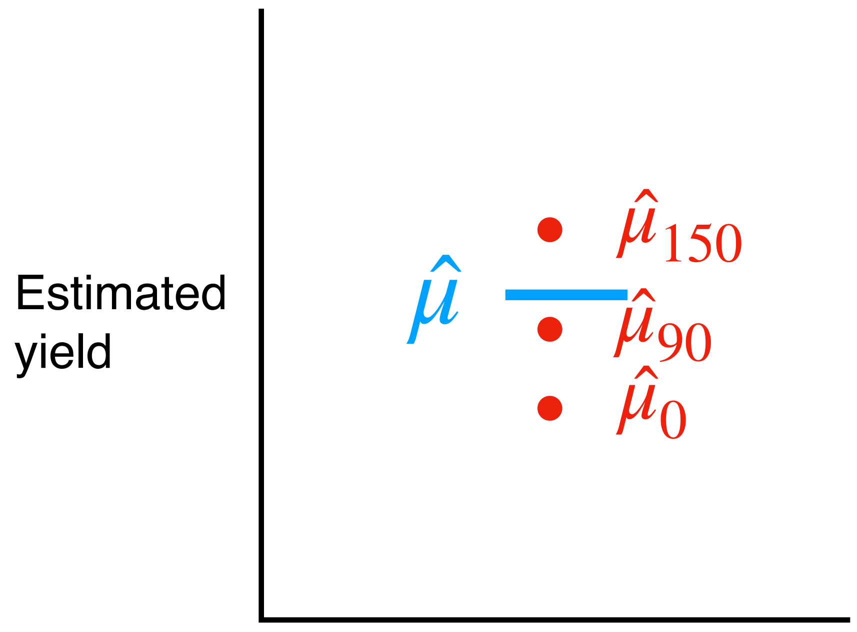
Estimate the **Variance** of **Treatment Means**

Normalize by the **Mean Squared Error**

Called an F-statistic

How much bigger is the estimated variance than what we'd expect if all means were the same?

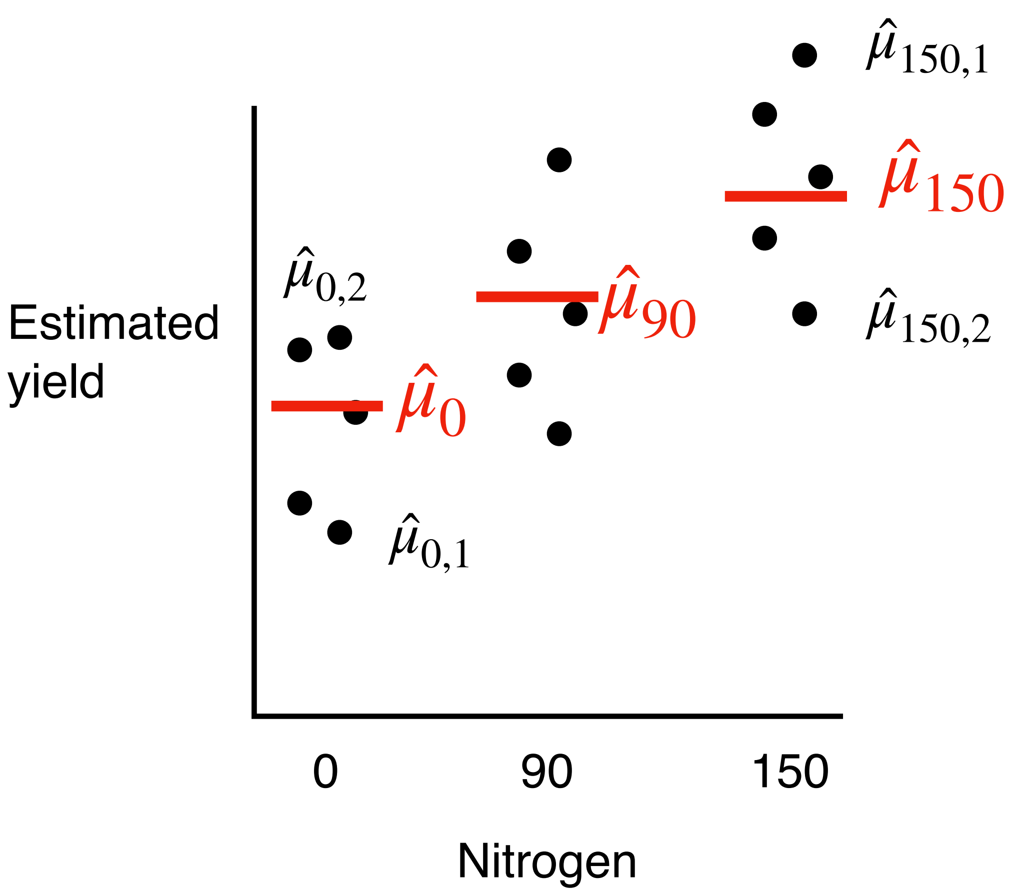
F-test



s_{μ}^2: Sample variance of estimated means

Degrees of Freedom: (k-1)

Estimate of: $\sigma_{\mu}^2 + \dfrac{\sigma_{\mu_i}^2 + \sigma_m^2}{n_i}$



s_{pooled}^2: Sample variance of EU estimates

Degrees of Freedom: k*$(n_i - 1)$

Estimate of: $\sigma_{\mu_i}^2 + \sigma_m^2$

We know n_i from our experimental design

If the Null Hypothesis is true and all treatment means are the same, $\sigma_{\mu}^2 = 0$

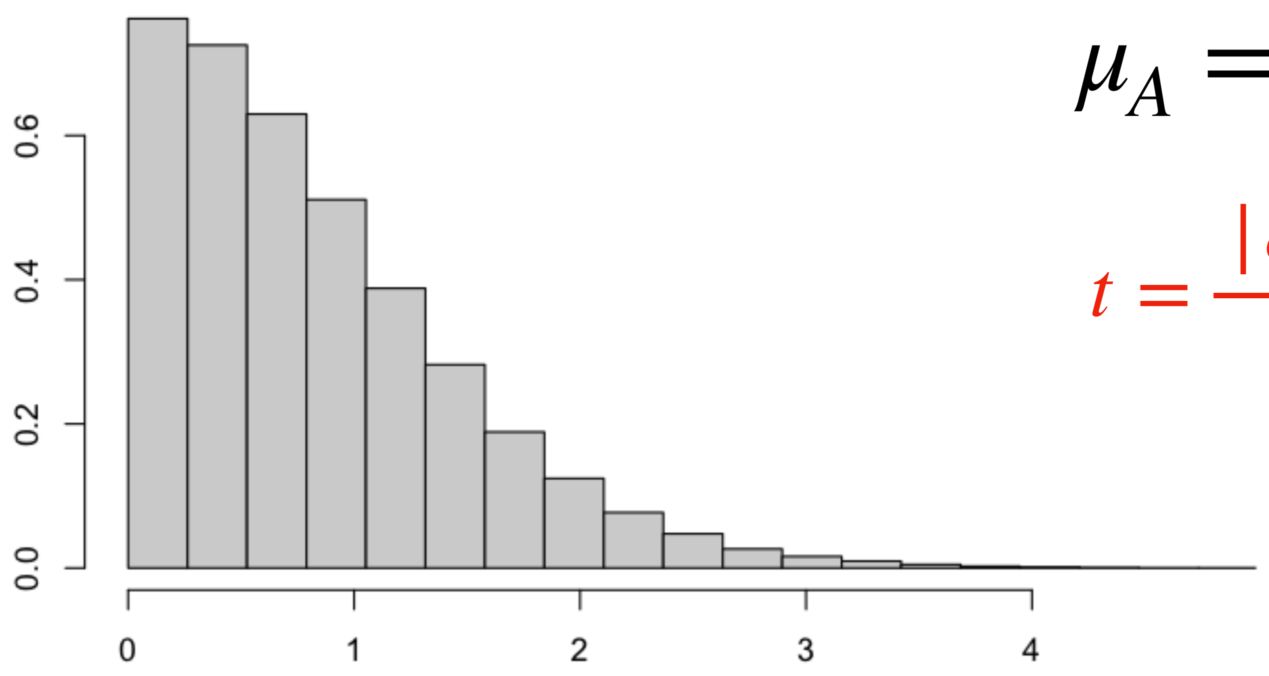
Then: $n_i \times s_{\mu}^2$ and s_{pooled}^2 both estimate the same thing

$$\frac{n_i s_{\mu}^2}{s_{pooled}^2}$$ should be ~1

If $\sigma_{\mu}^2 > 0$, $\dfrac{n_i s_{\mu}^2}{s_{pooled}^2} > 1$

T-test

Null Hypothesis: Both treatment means are equal



$$\mu_A = \mu_B \qquad \delta = 0$$

$$t = \frac{|\hat{\delta} - 0|}{\text{SED}}$$

Error in treatment effect estimate

Estimated average error

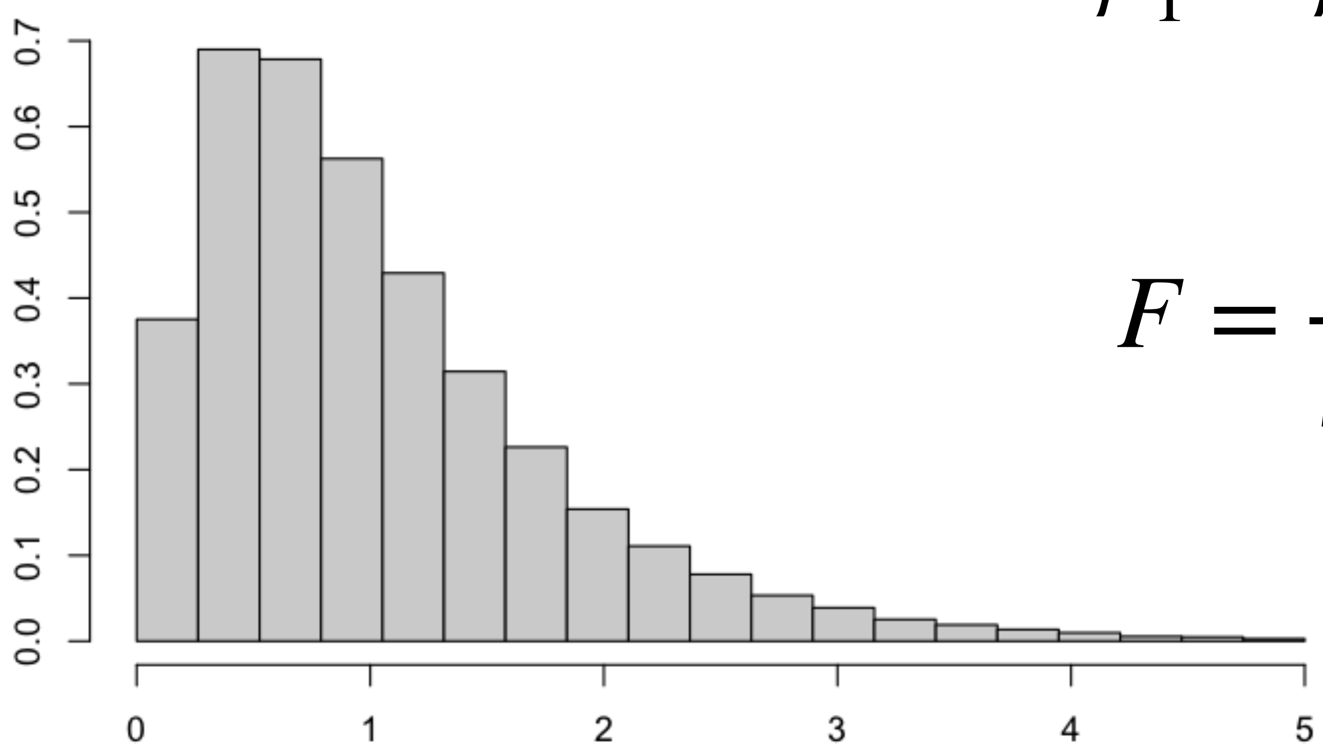
Big t implies H_0 is false

$$\text{SED} = \sqrt{\frac{s^2_{pooled}}{n_1} + \frac{s^2_{pooled}}{n_2}}$$

df from calculation of s^2_{pooled}

F-test

Null Hypothesis: All treatment means are equal



$$\mu_1 = \mu_2 = \ldots = \mu_k \qquad \sigma^2_\mu = 0$$

$$F = \frac{n_i s^2_\mu}{s^2_{pooled}}$$

~Error in treatment variance estimate

~Expected variance

Big F implies H_0 is false

df_2 from calculation of s^2_{pooled}

df_1 from calculation of s^2_μ

$F_{df1,df2}$

If k=2, $F = t^2$

The F-test is just a generalization of the T-test

Inherits all its power and its problems

Source	Df	SumSq	MeanSq	F-value	p-value
Treatment	dfT	SST	$\frac{SST}{dfT}$ = MST	F = $\frac{MST}{MSE}$	p
Error	dfE	SSE	$\frac{SSE}{dfE}$ = MSE		

$$F = \frac{n_i s_\mu^2}{s_{pooled}^2} = \frac{MST}{MSE}$$

$n_i s_\mu^2$ = MST = Sample Variance of estimated treatment means

$$= \frac{n_i * \text{Sum of Squared Deviations of Treatments}}{\text{\# treatments - 1}} = \frac{SST}{dfT}$$

s_{pooled}^2 = MSE = Sample Variance of estimated EUs

$$= \frac{\text{Sum of Squared Deviations of EUs}}{\text{\# independent EUs}} = \frac{SSE}{dfE}$$

p-value = pf(F,df1,df2,lower.tail=F) *no x2 here!