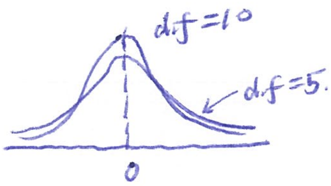
Introduction To statistics

* F distribution & F Test:

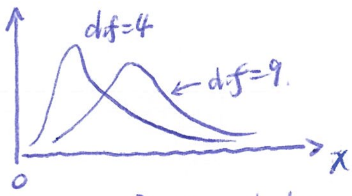
(I): Recall.



T-distribution.

can be +/-.

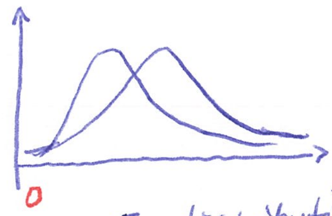
$$\frac{\bar{X}_n - u}{S/\sqrt{n}} \sim t_{(n-1)}$$



χ^2-distribution

only be +. $(X > 0)$

$$\frac{(n-1)S^2}{\sigma^2} = \frac{\Sigma(X_i - \bar{X}_n)^2}{\sigma^2} \sim \chi^2_{(n-1)}$$



F-distribution

can be +. only. $(X > 0)$

2 variances from 2 samples.

$$\frac{S_1^2}{S_2^2} \sim F_{(n_1-1,\, n_2-1)}$$

⇒: F-test: used to compare 2 variance.

Example: Medical research: whether the variance of heart rates of smokers is different from " " " " ... of Non-Smoker?

Randomly select 2 samples:

Smokers:	Non-Smokers
$n_1 = 26$	$n_2 = 18$
$S_1^2 = 36$	$S_2^2 = 10$.

⇒. Step 1: $\begin{cases} H_0: \ \sigma_1^2 = \sigma_2^2 \\ H_1: \ \sigma_1^2 \neq \sigma_2^2. \ (2\text{-sided}) \end{cases}$

Step 2: Find the critical value at $\alpha = 0.05$, Note it's 2-sided: from the F-distribution table, $\alpha/2 = 0.025$.

$d.f.N = 26 - 1 = 25$. ⇒. C.V. = 2.56.
$d.f.D = 18 - 1 = 17$.

Step 3: $F = \frac{S_1^2}{S_2^2} = \frac{36}{10} = 3.6$

Step 4: Since $3.6 > C.V. = 2.56$ ⇒ Reject H_0

Step 5: They are significant different at $\alpha = 0.05$.

(II): F-test: Analysis of Variance. (ANOVA).

Q: We can use Z-test, T-test to test the means of 2 groups.

What if we have 3 groups or more?

⇒ Any test to compare 3 means at one time (simutaneously)?

⇒ F-test can help us!

$$\begin{cases} H_0: \mu_1 = \mu_2 \cdots = \mu_k: & k\text{-groups.} \\ H_1: \text{At least one mean is different from the others.} \end{cases}$$

Example: Want to compare blood pressures of 3 groups:

	Medication	Exercise.	Diet.
reduction in blood pressure:	10	6	5
	12	8	9
	9	3	12
	15	0	8
	13	2	4.
Avg:	$\bar{X}_1 = 11.8$.	$\bar{X}_2 = 3.8$	$\bar{X}_3 = 7.6$
	$n_1 = 5$.	$n_2 = 5$	$n_3 = 5$
Within group variance ⇒	$S_1^2 = 5.7$.	$S_2^2 = 10.2$.	$S_3^2 = 10.3$.
Between group variance ⇒	$5(11.8-7.73)^2$	$5(3.8-7.73)^2$	$5(7.6-7.73)^2$.
	$\bar{X}_1 - \bar{\bar{X}}$		

Overall Average:

$$\bar{X}_{GM} = \frac{\Sigma X_i}{15} = 7.73$$

⇒ sum of within group variance. $(5-1)\times 5.7 + (5-1)\times 10.2 + (5-1)\cdot 10.3 = 104.80$

Avg. variance within group: $MS_W = \dfrac{104.80}{3(5-1)} = 8.73$

sum of between group variance: $5(11.8-7.73)^2 + 5(3.8-7.73)^2 + 5(7.6-7.73)^2 = 160.13$

Avg. variance between groups. $MS_B = \dfrac{160.13}{2} = 80.07$

⇒ $F = \dfrac{MS_B}{MS_W} = \dfrac{80.07}{8.73} = 9.17$.

Given $\alpha = 0.05$, d.f. N $= 3-1 = 2$ ⇒ C.V. $= 3.89$.

d.f. D $= 15-3 = 12$.

⇒ Since $9.17 > 3.89$, outside. reject H_0.

⇒ At least one group is different from the others.