

MSE-213

Probability and statistics for materials science

Lecture 10

Recap: ANOVA / F-test for groups/discrete factors

Source of Variation	Degrees of Freedom, ν	Sum of Squares, SS	Mean square, MS	Fisher statistic, F_{MEASURED}
Between Groups/Factors	$I-1$	SS_B	$SS_B/(I-1)$	MS_B/MS_E
Error within Group	N_T-I	SS_E	$SS_E/(N_T-I)$	
Total	N_T-1	SS_T		

Quantiles de la loi F_{ν_1, ν_2} de Fisher

	$\nu_1 = 1$	2	3	4	5	6	7	8	10	12	24	∞
$\nu_2 = 1$	161,4	199,5	215,7	224,6	230,2	234,0	236,8	238,9	241,9	243,9	249,1	254,3
2	18,51	19,00	19,16	19,25	19,30	19,33	19,35	19,37	19,40	19,41	19,45	19,50
3	10,13	9,552	9,277	9,117	9,013	8,941	8,887	8,845	8,786	8,745	8,639	8,526
4	7,709	6,944	6,591	6,388	6,256	6,163	6,094	6,041	5,964	5,912	5,774	5,628
5	6,608	5,786	5,409	5,192	5,050	4,950	4,876	4,818	4,735	4,678	4,527	4,365
6	5,987	5,143	4,757	4,534	4,387	4,284	4,207	4,147	4,060	4,000	3,841	3,669
7	5,591	4,737	4,347	4,120	3,972	3,866	3,787	3,726	3,637	3,575	3,410	3,230
8	5,318	4,459	4,066	3,838	3,687	3,581	3,500	3,438	3,347	3,284	3,115	2,928
9	5,117	4,256	3,863	3,633	3,482	3,374	3,293	3,230	3,137	3,073	2,900	2,707
10	4,965	4,103	3,708	3,478	3,326	3,217	3,135	3,072	2,978	2,913	2,737	2,538
11	4,844	3,982	3,587	3,357	3,204	3,095	3,012	2,948	2,854	2,788	2,609	2,404
12	4,747	3,885	3,490	3,259	3,106	2,996	2,913	2,849	2,753	2,687	2,505	2,296
13	4,667	3,806	3,411	3,179	3,025	2,915	2,832	2,767	2,671	2,604	2,420	2,206
14	4,600	3,739	3,344	3,112	2,958	2,848	2,764	2,699	2,602	2,534	2,349	2,131
15	4,543	3,682	3,287	3,056	2,901	2,790	2,707	2,641	2,544	2,475	2,288	2,066
16	4,494	3,634	3,239	3,007	2,852	2,741	2,657	2,591	2,494	2,425	2,235	2,010
17	4,451	3,592	3,197	2,965	2,810	2,699	2,614	2,548	2,450	2,381	2,190	1,960
18	4,414	3,555	3,160	2,928	2,773	2,661	2,577	2,510	2,412	2,342	2,150	1,917
19	4,381	3,522	3,127	2,895	2,740	2,628	2,544	2,477	2,378	2,308	2,114	1,878
20	4,351	3,493	3,098	2,866	2,711	2,599	2,514	2,447	2,348	2,278	2,082	1,843
21	4,325	3,467	3,072	2,840	2,685	2,573	2,488	2,420	2,321	2,250	2,054	1,812
22	4,301	3,443	3,049	2,817	2,661	2,549	2,464	2,397	2,297	2,226	2,028	1,783
23	4,279	3,422	3,028	2,796	2,640	2,528	2,442	2,375	2,275	2,204	2,005	1,757
24	4,260	3,403	3,009	2,776	2,621	2,508	2,423	2,355	2,255	2,183	1,984	1,733
25	4,242	3,385	2,991	2,759	2,603	2,490	2,405	2,337	2,236	2,165	1,964	1,711
26	4,225	3,369	2,975	2,743	2,587	2,474	2,388	2,321	2,220	2,148	1,946	1,691
27	4,210	3,354	2,960	2,728	2,572	2,459	2,373	2,305	2,204	2,132	1,930	1,672
28	4,196	3,340	2,947	2,714	2,558	2,445	2,359	2,291	2,190	2,118	1,915	1,654
29	4,183	3,328	2,934	2,701	2,545	2,432	2,346	2,278	2,177	2,104	1,901	1,638
30	4,171	3,316	2,922	2,690	2,534	2,421	2,334	2,266	2,165	2,092	1,887	1,622
32	4,149	3,295	2,901	2,668	2,512	2,399	2,313	2,244	2,142	2,070	1,864	1,594
34	4,130	3,276	2,883	2,650	2,494	2,380	2,294	2,225	2,123	2,050	1,843	1,569
36	4,113	3,259	2,866	2,634	2,477	2,364	2,277	2,209	2,106	2,033	1,824	1,547
38	4,098	3,245	2,852	2,619	2,463	2,349	2,262	2,194	2,091	2,017	1,808	1,527
40	4,085	3,232	2,839	2,606	2,449	2,336	2,249	2,180	2,077	2,003	1,793	1,509
60	4,001	3,150	2,758	2,525	2,368	2,254	2,167	2,097	1,993	1,917	1,700	1,389
120	3,920	3,072	2,680	2,447	2,290	2,175	2,087	2,016	1,910	1,834	1,608	1,254
∞	3,841	2,996	2,605	2,372	2,214	2,099	2,010	1,938	1,831	1,752	1,517	1,000

TABLE 5 – Les 95%-quantiles, $qF_{\nu_1, \nu_2}(95\%)$, des distributions F_{ν_1, ν_2} .

Quantiles de la loi

	$\nu_1 = 1$	2	3	4	5
$\nu_2 = 1$	161,4	199,5	215,7	224,6	230,2
2	18,51	19,00	19,16	19,25	19,30
3	10,13	9,552	9,277	9,117	9,013
4	7,709	6,944	6,591	6,388	6,256
5	6,608	5,786	5,409	5,192	5,050
6	5,987	5,143	4,757	4,534	4,387
7	5,591	4,737	4,347	4,120	3,972
8	5,318	4,459	4,066	3,838	3,687
9	5,117	4,256	3,863	3,633	3,482
10	4,965	4,103	3,708	3,478	3,326
11	4,844	3,982	3,587	3,357	3,204

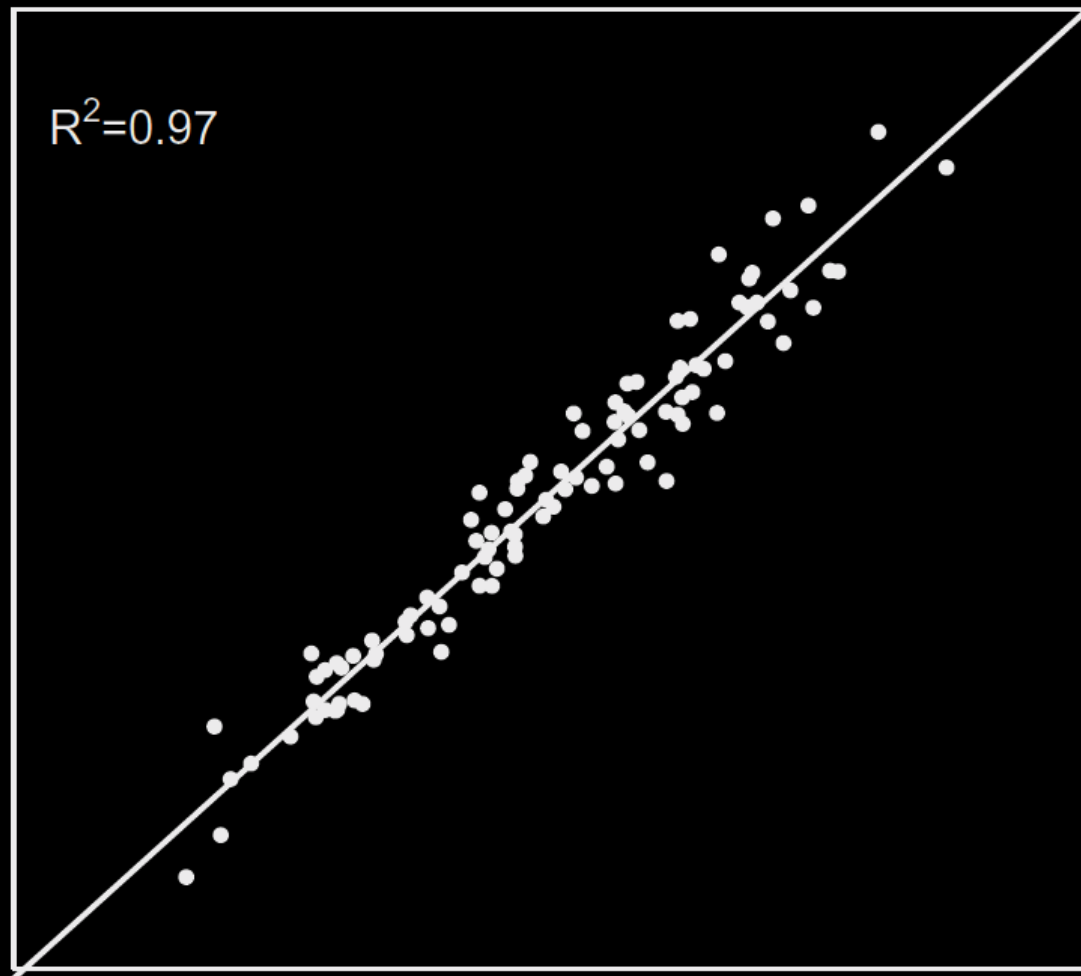
Linear Regression with one slope

ANOVA / F-test for linear regression with one slope

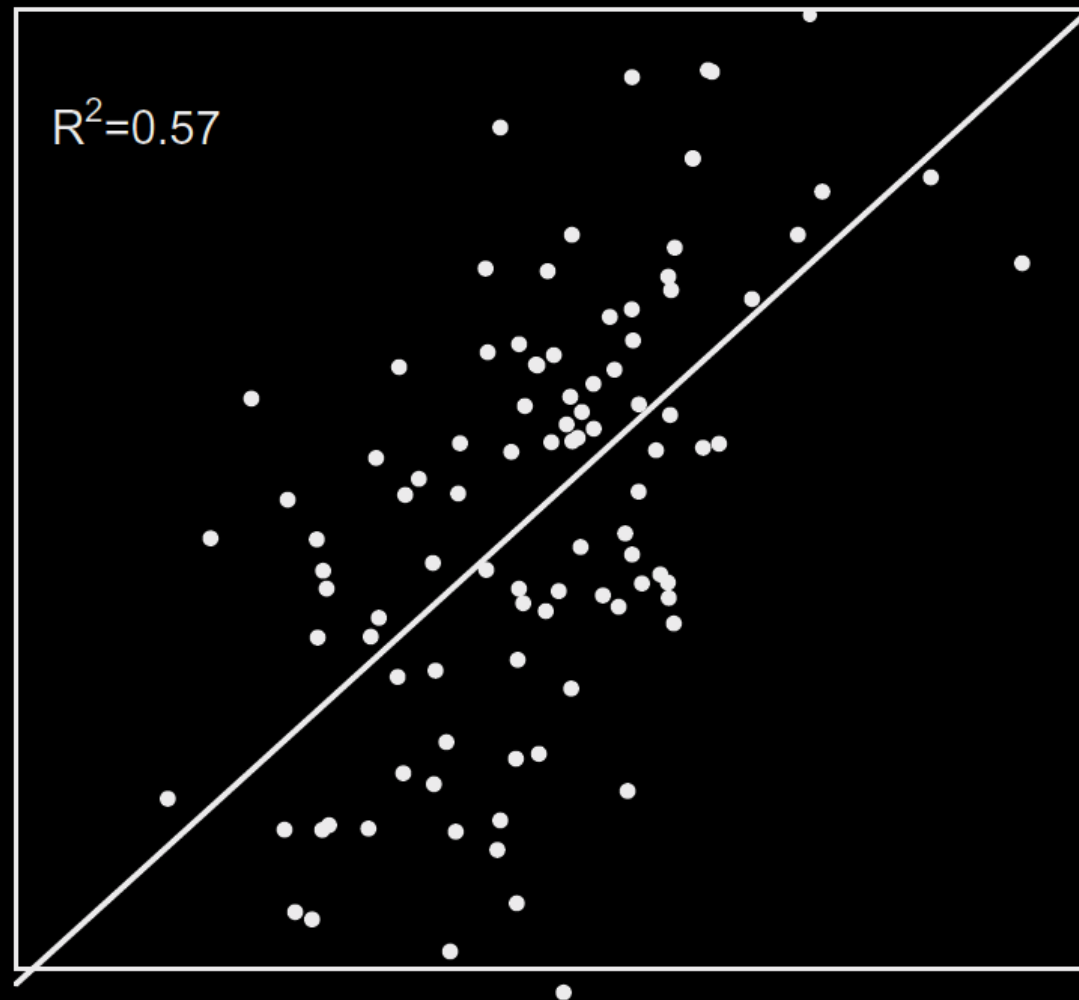
Source of Variation	Degrees of Freedom, ν	Sum of Squares, SS	Mean square, MS	Fisher statistic, F_{MEASURED}
Model (with 1 slope)	1	SS_M	$SS_M/1$	MS_M/MS_E
Error within Residuals	N_T-2	SS_E	$SS_E/(N_T-2)$	
Total	N_T-1	SS_T		

SS_M sometimes called SS_R for "REGRESSION"

Linear Regression – The R^2 “goodness of fit”



(a) Strong linear relationship



(b) Weak linear relationship

BUT: All R^2 are 0.82 here

