

分散の加法性を視覚的に理解する（その4）

Sampo Suzuki, CC 4.0 BY-NC-SA

2021-06-06

正規分布とは異なる分布でも分散の加法性が成り立つことを確認します。

t 分布の場合

自由度 $df = df = \text{Inf}$ の t 分布の分散を以下の関数で計算します。

```
1 ft <- function(i = NA, n = 5000000) {
2   x <- rt(n = n, df = 5)
3   y <- rt(n = n, df = Inf)
4   df <- data.frame(no = i,
5                     var.x = var(x), var.y = var(y),
6                     var.xy = var(x + y), var.sum = var(x) + var(y),
7                     cov2 = cov(x, y) * 2)
8   df <- cor.test(x, y) %>% broom::tidy() %>% dplyr::bind_cols(df)
9   return(df)
10 }
```

加法 1: $var.xy = var(x + y)$, 加法 2: $var.sum = var(x) + var(y)$

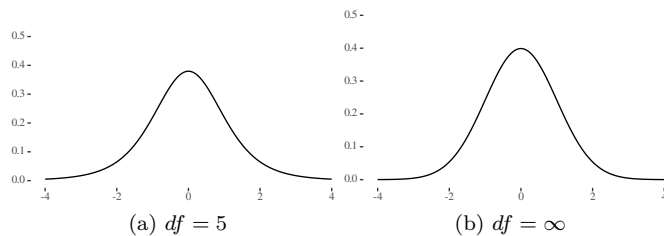


Figure 1: 分散を加算する二種類の t 分布

Table 1: 計算結果

No	相関係数	p 値	標本 x	標本 y	加法 1	加法 2	差異	加法 1/加法 2	cov2
1	-0.0003400	0.4471338	1.671315	1.0006527	2.671088	2.671968	-0.0008793	0.9996709	-0.0008793
2	-0.0001132	0.8001465	1.665842	1.0002102	2.665760	2.666052	-0.0002923	0.9998904	-0.0002923
3	0.0001970	0.6595409	1.666911	1.0003627	2.667783	2.667274	0.0005088	1.0001908	0.0005088
4	0.0002207	0.6216512	1.667868	1.0005398	2.668978	2.668408	0.0005702	1.0002137	0.0005702
5	0.0003219	0.4716171	1.665999	1.0010391	2.667870	2.667038	0.0008315	1.0003118	0.0008315
6	-0.0001162	0.7950144	1.668415	0.9996618	2.667776	2.668076	-0.0003001	0.9998875	-0.0003001
7	0.0002744	0.5394774	1.667536	1.0018700	2.670115	2.669405	0.0007094	1.0002657	0.0007094
8	0.0005338	0.2326176	1.665655	0.9998358	2.666869	2.665491	0.0013778	1.0005169	0.0013778
9	0.0005526	0.2166069	1.668738	0.9999556	2.670121	2.668693	0.0014276	1.0005349	0.0014276
10	-0.0009500	0.0336438	1.670255	1.0000601	2.667859	2.670315	-0.0024557	0.9990804	-0.0024557
11	-0.0004474	0.3170860	1.668672	0.9997065	2.667223	2.668379	-0.0011558	0.9995669	-0.0011558
12	0.0003121	0.4852412	1.669184	0.9993104	2.669301	2.668494	0.0008062	1.0003021	0.0008062
13	-0.0007914	0.0767747	1.663218	1.0007921	2.661968	2.664010	-0.0020422	0.9992334	-0.0020422
14	0.0001397	0.7547932	1.665388	1.0000734	2.665822	2.665461	0.0003605	1.0001353	0.0003605
15	-0.0000405	0.9278549	1.668708	0.9992771	2.667881	2.667985	-0.0001046	0.9999608	-0.0001046
16	-0.0005011	0.2625260	1.663565	1.0007718	2.663044	2.664337	-0.0012931	0.9995147	-0.0012931
17	-0.0006516	0.1450992	1.665783	0.9992144	2.663316	2.664997	-0.0016814	0.9993691	-0.0016814
18	0.0003550	0.4272851	1.670817	0.9996750	2.671410	2.670492	0.0009176	1.0003436	0.0009176
19	0.0005604	0.2102049	1.667328	1.0008842	2.669660	2.668212	0.0014478	1.0005426	0.0014478
20	-0.0003488	0.4353974	1.664794	1.0002748	2.664169	2.665069	-0.0009003	0.9996622	-0.0009003

χ^2 分布の場合

自由度 $df = 1$ と $df = 3$ の χ^2 分布の分散を以下の関数で計算します。

```

1 fchisq <- function(i = NA, n = 5000000) {
2   x <- rchisq(n = n, df = 1)
3   y <- rchisq(n = n, df = 3)
4   df <- data.frame(no = i,
5                     var.x = var(x), var.y = var(y),
6                     var.xy = var(x + y), var.sum = var(x) + var(y),
7                     cov2 = cov(x, y) * 2)
8   df <- cor.test(x, y) %>% broom::tidy() %>% dplyr::bind_cols(df)
9   return(df)
10 }
```

加法1: $var.xy = var(x + y)$, 加法2: $var.sum = var(x) + var(y)$

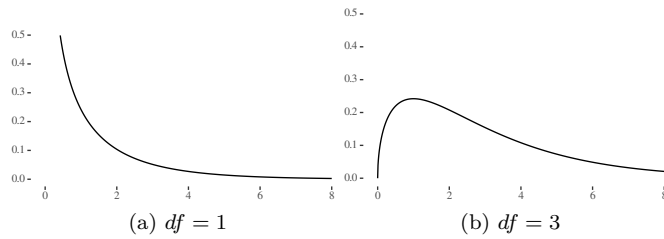


Figure 2: 分散を加算する二種類の χ^2 分布

Table 2: 計算結果

No	相関係数	p 値	標本 x	標本 y	加法 1	加法 2	差異	加法 1/加法 2	cov2
1	0.0004956	0.2677546	2.004440	6.012149	8.020031	8.016590	0.0034411	1.0004292	0.0034411
2	0.0000173	0.9691781	1.997486	5.998751	7.996357	7.996237	0.0001196	1.0000150	0.0001196
3	0.0000429	0.9236517	1.999706	6.012871	8.012874	8.012577	0.0002972	1.0000371	0.0002972
4	0.0000098	0.9825511	1.997608	5.997723	7.995399	7.995331	0.0000677	1.0000085	0.0000677
5	0.0003178	0.4772687	1.999837	5.988250	7.990287	7.988087	0.0021998	1.0002754	0.0021998
6	-0.0007314	0.1019382	1.995278	5.993983	7.984202	7.989261	-0.0050590	0.9993668	-0.0050590
7	-0.0002029	0.6500457	1.999653	5.996046	7.994294	7.995699	-0.0014051	0.9998243	-0.0014051
8	-0.0002943	0.5105458	1.998821	6.004295	8.001078	8.003117	-0.0020388	0.9997452	-0.0020388
9	0.0000522	0.9070285	1.996728	5.992009	7.989099	7.988738	0.0003613	1.0000452	0.0003613
10	-0.0001619	0.7174116	1.997968	5.998410	7.995258	7.996379	-0.0011207	0.9998599	-0.0011207
11	0.0003936	0.3787565	1.998269	5.998312	7.999307	7.996581	0.0027256	1.0003408	0.0027256
12	0.0001840	0.6807879	1.997278	5.997512	7.996063	7.994790	0.0012735	1.0001593	0.0012735
13	-0.0003081	0.4909276	2.007474	6.003329	8.008665	8.010804	-0.0021389	0.9997330	-0.0021389
14	-0.0001352	0.7623968	2.006404	6.001392	8.006858	8.007796	-0.0009384	0.9998828	-0.0009384
15	0.0000724	0.8714218	2.001221	5.999281	8.001004	8.000502	0.0005016	1.0000627	0.0005016
16	-0.0005146	0.2498236	1.997352	6.002466	7.996254	7.999818	-0.0035639	0.9995545	-0.0035639
17	0.0002133	0.6333729	1.997627	6.003839	8.002944	8.001467	0.0014775	1.0001847	0.0014775
18	-0.0004308	0.3354494	2.001473	6.002582	8.001068	8.004055	-0.0029861	0.9996269	-0.0029861
19	0.0002259	0.6134288	1.995501	5.998105	7.995169	7.993606	0.0015633	1.0001956	0.0015633
20	-0.0006592	0.1404616	1.997356	6.002748	7.995539	8.000104	-0.0045653	0.9994293	-0.0045653

F 分布の場合

自由度 $df_1 = 3, df_2 = 6$ と $df_1 = 9, df_2 = 3$ の F 分布の分散を以下の関数で計算します。

```
1 ff <- function(i = NA, n = 5000000) {
2   x <- rf(n = n, df1 = 3, df2 = 6)
3   y <- rf(n = n, df1 = 9, df2 = 3)
4   df <- data.frame(no = i,
5                     var.x = var(x), var.y = var(y),
6                     var.xy = var(x + y), var.sum = var(x) + var(y),
7                     cov2 = cov(x, y) * 2)
8   df <- cor.test(x, y) %>% broom::tidy() %>% dplyr::bind_cols(df)
9   return(df)
10 }
```

加法 1: $var.xy = var(x + y)$, 加法 2: $var.sum = var(x) + var(y)$

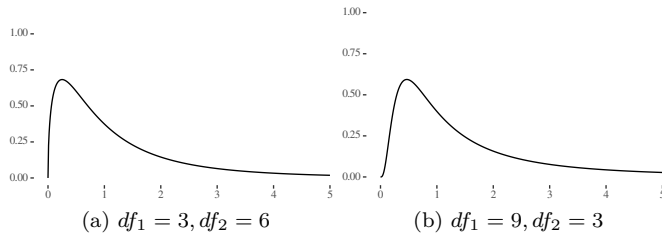


Figure 3: 分散を加算する二種類の F 分布

Table 3: 計算結果

No	相関係数	p 値	標本 x	標本 y	加法 1	加法 2	差異	加法 1/加法 2	cov2
1	-0.0005088	0.2552622	5.211667	2187.4152	2192.5182	2192.6269	-0.1086455	0.9999504	-0.1086455
2	-0.0000916	0.8376373	5.108454	1445.2154	1450.3081	1450.3239	-0.0157483	0.9999891	-0.0157483
3	-0.0002445	0.5845768	5.199387	2018.1786	2023.3279	2023.3780	-0.0500910	0.9999752	-0.0500910
4	0.0002644	0.5544113	5.122349	469.2528	474.4011	474.3752	0.0259233	1.0000546	0.0259233
5	-0.0001986	0.6569190	5.215827	1280.2691	1285.4525	1285.4849	-0.0324644	0.9999747	-0.0324644
6	-0.0002592	0.5622175	5.434277	1433.5962	1438.9847	1439.0304	-0.0457532	0.9999682	-0.0457532
7	-0.0001976	0.6585898	5.176098	48599.2364	48604.2143	48604.4125	-0.1982197	0.9999959	-0.1982197
8	0.0001127	0.8009720	5.074826	2330.8257	2335.9251	2335.9006	0.0245225	1.0000105	0.0245225
9	-0.0002067	0.6439203	5.159144	1372.0367	1377.1611	1377.1959	-0.0347832	0.9999747	-0.0347832
10	-0.0003368	0.4514110	5.182321	744.9437	750.0842	750.1260	-0.0418506	0.9999442	-0.0418506
11	0.0001890	0.6725413	5.312891	748.5687	753.9055	753.8816	0.0238408	1.0000316	0.0238408
12	0.0001013	0.8207210	5.098162	1298.4698	1303.5845	1303.5680	0.0164916	1.0000127	0.0164916
13	0.0002387	0.5935867	5.178266	1903.9322	1909.1579	1909.1105	0.0473932	1.0000248	0.0473932
14	0.0005441	0.2237773	5.220821	749.8635	755.1525	755.0844	0.0680822	1.0000902	0.0680822
15	0.0002153	0.6302591	5.207314	1537.7464	1542.9922	1542.9537	0.0385271	1.0000250	0.0385271
16	0.0003802	0.3951848	5.181432	974.5329	979.7684	979.7144	0.0540401	1.0000552	0.0540401
17	0.0003537	0.4289961	5.123908	648.6608	653.8255	653.7847	0.0407833	1.0000624	0.0407833
18	-0.0003690	0.4093272	5.272903	813.8613	819.0858	819.1342	-0.0483437	0.9999410	-0.0483437
19	0.0002439	0.5854696	5.635600	933.4921	939.1631	939.1277	0.0353831	1.0000377	0.0353831
20	0.0005904	0.1867975	5.529007	1510.8539	1516.4908	1516.3829	0.1079172	1.0000712	0.1079172

t分布と χ^2 分布の場合

自由度 $df = 5$ の t 分布と $df = 3$ の χ^2 分布の分散を以下の関数で計算します。

```
1 ftchisq <- function(i = NA, n = 5000000) {
2   x <- rt(n = n, df = 5)
3   y <- rchisq(n = n, df = 3)
4   df <- data.frame(no = i,
5                     var.x = var(x), var.y = var(y),
6                     var.xy = var(x + y), var.sum = var(x) + var(y),
7                     cov2 = cov(x, y) * 2)
8   df <- cor.test(x, y) %>% broom::tidy() %>% dplyr::bind_cols(df)
9   return(df)
10 }
```

加法1: $var.xy = var(x + y)$, 加法2: $var.sum = var(x) + var(y)$

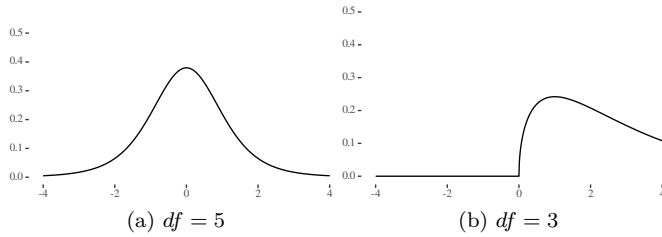


Figure 4: 分散を加算する二種類の χ^2 分布

Table 4: 計算結果

No	相関係数	p 値	標本 x	標本 y	加法 1	加法 2	差異	加法 1/加法 2	cov2
1	-0.0001782	0.6903304	1.664412	5.999207	7.662493	7.663619	-0.0011260	0.9998531	-0.0011260
2	-0.0002724	0.5425179	1.664679	6.004114	7.667071	7.668793	-0.0017221	0.9997754	-0.0017221
3	-0.0001043	0.8155900	1.666044	5.991276	7.656660	7.657319	-0.0006590	0.9999139	-0.0006590
4	-0.0007060	0.1144232	1.666784	6.009416	7.671731	7.676200	-0.0044687	0.9994179	-0.0044687
5	0.0005852	0.1906797	1.664959	5.998756	7.667414	7.663715	0.0036989	1.0004827	0.0036989
6	-0.0008875	0.0472098	1.666058	5.995425	7.655873	7.661483	-0.0056096	0.9992678	-0.0056096
7	0.0000842	0.8506780	1.667544	5.999345	7.667422	7.666890	0.0005326	1.0000695	0.0005326
8	0.0001750	0.6955518	1.668354	6.008008	7.677471	7.676363	0.0011082	1.0001444	0.0011082
9	-0.0000253	0.9549389	1.669423	6.001814	7.671077	7.671237	-0.0001600	0.9999791	-0.0001600
10	0.0001198	0.7887721	1.668855	6.006666	7.676280	7.675521	0.0007587	1.0000988	0.0007587
11	0.0002645	0.5542274	1.668434	6.004580	7.674688	7.673014	0.0016744	1.0002182	0.0016744
12	-0.0003488	0.4354460	1.666775	5.999497	7.664066	7.666271	-0.0022059	0.9997123	-0.0022059
13	0.0007602	0.0891628	1.663094	6.006085	7.673984	7.669179	0.0048051	1.0006266	0.0048051
14	-0.0000040	0.9928889	1.664894	5.995082	7.659951	7.659976	-0.0000252	0.9999967	-0.0000252
15	0.0005534	0.2159302	1.669772	6.003873	7.677149	7.673645	0.0035043	1.0004567	0.0035043
16	0.0000438	0.9219608	1.666002	5.998464	7.664743	7.664466	0.0002770	1.0000361	0.0002770
17	0.0002501	0.5760000	1.665924	6.003260	7.670766	7.669184	0.0015818	1.0002063	0.0015818
18	-0.0004545	0.3094799	1.665073	5.992650	7.654852	7.657724	-0.0028714	0.9996250	-0.0028714
19	-0.0002343	0.6003523	1.664831	6.006462	7.669810	7.671292	-0.0014818	0.9998068	-0.0014818
20	0.0001036	0.8167538	1.670891	6.008467	7.680015	7.679358	0.0006567	1.0000855	0.0006567