

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

2021-06-05

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

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

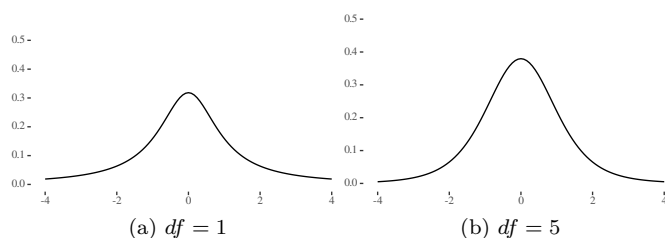


Figure 1: 分散を加算する二種類の t 分布

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	-0.0001503	0.7367569	3756872	1.668392	939218.2	939218.4	-0.1881836	0.9999998	-0.1881836
2	-0.00009528	0.0331366	5921163	1.665188	1480289.7	1480291.1	-1.4958439	0.9999990	-1.4958439
3	-0.0001065	0.8117592	3170503	1.665145	792626.1	792626.2	-0.1223598	0.9999998	-0.1223598
4	0.0001363	0.7605916	234538469	1.668358	58634619.0	58634617.6	1.3477671	1.0000000	1.3477671
5	-0.0003846	0.3898150	21328947	1.667404	5332236.1	5332237.2	-1.1467396	0.9999998	-1.1467396
6	0.0000707	0.8744723	2824118	1.670452	706030.1	706030.0	0.0767267	1.0000001	0.0767267
7	-0.0003141	0.4824975	31822527	1.667388	7955631.0	7955632.1	-1.1438998	0.9999999	-1.1438998
8	-0.0001649	0.7124014	29039432	1.665014	7259857.9	7259858.5	-0.5731684	0.9999999	-0.5731684
9	-0.0001918	0.6680222	16324840	1.667192	4081210.0	4081210.5	-0.5002902	0.9999999	-0.5002902
10	-0.0006548	0.1431528	1649851	1.668310	412462.5	412463.0	-0.5431646	0.9999987	-0.5431646
11	-0.0004414	0.3235916	39759746	1.667619	9939935.2	9939937.0	-1.7972904	0.9999998	-1.7972904
12	0.0005128	0.2515537	163643080	1.664719	40910774.7	40910770.5	4.2316546	1.0000001	4.2316546
13	-0.0004253	0.3416042	11220099	1.666779	2805024.3	2805025.2	-0.9196087	0.9999997	-0.9196087
14	-0.0001211	0.7865968	61321544	1.666002	15330385.8	15330386.4	-0.6118848	1.0000000	-0.6118848
15	-0.0000186	0.9668654	431531	1.665694	107883.2	107883.2	-0.0078751	0.9999999	-0.0078751
16	0.0003024	0.4988689	7363469	1.664964	1840868.3	1840867.8	0.5294797	1.0000003	0.5294797
17	-0.0000513	0.9087598	611103876	1.666259	152775968.7	152775969.5	-0.8177302	1.0000000	-0.8177302
18	-0.0002512	0.5743927	1643759	1.666420	410939.9	410940.1	-0.2078351	0.9999995	-0.2078351
19	0.0000852	0.8489663	2205786	1.668551	551446.9	551446.9	0.0816938	1.0000001	0.0816938
20	0.0000339	0.9395056	471213880	1.669165	117803471.0	117803470.5	0.4759216	1.0000000	0.4759216

Table 1: 計算結果

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, num = 2) {
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) / num), var.sum = (var(x / num) + var(y / num)),
7                     cov2 = cov(x / num, y / num) * 2)
8   df <- cor.test(x, y) %>% broom::tidy() %>% dplyr::bind_cols(df)
9   return(df)
10 }
```

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

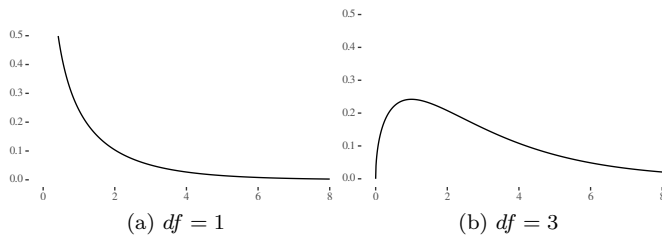


Figure 2: 分散を加算する二種類の χ^2 分布

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	-0.0001503	0.7367569	3756872	1.668392	939218.2	939218.4	-0.1881836	0.9999998	-0.1881836
2	-0.0009528	0.0331366	5921163	1.665188	1480289.7	1480291.1	-1.4958439	0.9999990	-1.4958439
3	-0.0001065	0.8117592	3170503	1.665145	792626.1	792626.2	-0.1223598	0.9999998	-0.1223598
4	0.0001363	0.7605916	234538469	1.668358	58634619.0	58634617.6	1.3477671	1.0000000	1.3477671
5	-0.0003846	0.3898150	21328947	1.667404	5332236.1	5332237.2	-1.1467396	0.9999998	-1.1467396
6	0.0000707	0.8744723	2824118	1.670452	706030.1	706030.0	0.0767267	1.0000001	0.0767267
7	-0.0003141	0.4824975	31822527	1.667388	7955631.0	7955632.1	-1.1438998	0.9999999	-1.1438998
8	-0.0001649	0.7124014	29039432	1.665014	7259857.9	7259858.5	-0.5731684	0.9999999	-0.5731684
9	-0.0001918	0.6680222	16324840	1.667192	4081210.0	4081210.5	-0.5002902	0.9999999	-0.5002902
10	-0.0006548	0.1431528	1649851	1.668310	412462.5	412463.0	-0.5431646	0.9999987	-0.5431646
11	-0.0004414	0.3235916	39759746	1.667619	9939935.2	9939937.0	-1.7972904	0.9999998	-1.7972904
12	0.0005128	0.2515537	163643080	1.664719	40910774.7	40910770.5	4.2316546	1.0000001	4.2316546
13	-0.0004253	0.3416042	11220099	1.666779	2805024.3	2805025.2	-0.9196087	0.9999997	-0.9196087
14	-0.0001211	0.7865968	61321544	1.666002	15330385.8	15330386.4	-0.6118848	1.0000000	-0.6118848
15	-0.0000186	0.9668654	431531	1.665694	107883.2	107883.2	-0.0078751	0.9999999	-0.0078751
16	0.0003024	0.4988689	7363469	1.664964	1840868.3	1840867.8	0.5294797	1.0000003	0.5294797
17	-0.0000513	0.9087598	611103876	1.666259	152775968.7	152775969.5	-0.8177302	1.0000000	-0.8177302
18	-0.0002512	0.5743927	1643759	1.666420	410939.9	410940.1	-0.2078351	0.9999995	-0.2078351
19	0.0000852	0.8489663	2205786	1.668551	551446.9	551446.9	0.0816938	1.0000001	0.0816938
20	0.0000339	0.9395056	471213880	1.669165	117803471.0	117803470.5	0.4759216	1.0000000	0.4759216

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, num = 2) {
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) / num), var.sum = (var(x / num) + var(y / num)),
7                     cov2 = cov(x / num, y / num) * 2)
8   df <- cor.test(x, y) %>% broom::tidy() %>% dplyr::bind_cols(df)
9   return(df)
10 }
```

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

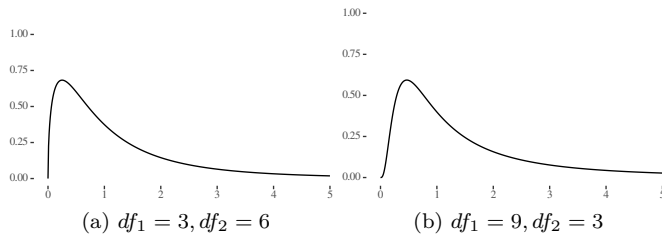


Figure 3: 分散を加算する二種類の F 分布

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	0.0002393	0.5926248	5.283749	9299.3652	2326.1888	2326.1622	0.0265195	1.0000114	0.0265195
2	-0.0002609	0.5596816	5.140466	6937.4608	1735.6257	1735.6503	-0.0246314	0.9999858	-0.0246314
3	-0.0005290	0.2368538	5.126077	1723.8906	432.2293	432.2542	-0.0248643	0.9999425	-0.0248643
4	0.0001149	0.7971652	5.262574	1077.7077	270.7469	270.7426	0.0043281	1.0000160	0.0043281
5	0.0001032	0.8175539	5.159225	1275.7443	320.2301	320.2259	0.0041850	1.0000131	0.0041850
6	0.0003099	0.4883332	5.203402	1559.4609	391.1800	391.1661	0.0139581	1.0000357	0.0139581
7	0.0000123	0.9780831	5.103390	925.8786	232.7459	232.7455	0.0004223	1.0000018	0.0004223
8	0.0002263	0.6127805	5.339981	964.3089	242.4203	242.4122	0.0081209	1.0000335	0.0081209
9	0.0007682	0.0858299	5.211663	1180.5351	296.4668	296.4367	0.0301293	1.0001016	0.0301293
10	-0.0003568	0.4249735	5.127960	864.1476	217.3070	217.3189	-0.0118757	0.9999454	-0.0118757
11	-0.0001285	0.7738059	5.161446	3846.1338	962.8147	962.8238	-0.0090546	0.9999906	-0.0090546
12	0.0000877	0.8445989	5.157615	1642.7133	411.9718	411.9677	0.0040344	1.0000098	0.0040344
13	-0.0002277	0.6106733	5.171313	221321.9773	55331.6654	55331.7872	-0.1217900	0.9999978	-0.1217900
14	0.0000624	0.8890394	5.350175	1236.1945	310.3887	310.3862	0.0025372	1.0000082	0.0025372
15	0.0001417	0.7512863	5.163609	2524.0812	632.3193	632.3112	0.0080909	1.0000128	0.0080909
16	0.0004361	0.3294718	5.346270	2559.8011	641.3124	641.2868	0.0255093	1.0000398	0.0255093
17	0.0001128	0.8009107	5.272119	12234.8436	3060.0432	3060.0289	0.0143208	1.0000047	0.0143208
18	0.0003906	0.3824324	5.211134	3601.8563	901.7936	901.7669	0.0267571	1.0000297	0.0267571
19	-0.0002745	0.5393874	5.508024	1170.0237	293.8719	293.8829	-0.0110170	0.9999625	-0.0110170
20	-0.0004689	0.2944065	5.239619	6599.0153	1651.0201	1651.0637	-0.0435958	0.9999736	-0.0435958