

9-2 이론적 확률분포(연속형) 4



9-2-1. t-분포

t-분포

$Z \sim N(0,1), V \sim \chi^2(v)$ 이고 Z 와 V 가 독립이면, $T = \frac{Z}{\sqrt{\frac{V}{v}}} \sim t(v)$ 이다.

$$t = \frac{\bar{X} - \mu}{\frac{s}{\sqrt{n}}} \sim t_{n-1}$$

자유도가 무한히 커지면 **t**-분포는 표준정규분포에 접근한다.

$(X_1, X_2, \dots, X_n)$ 을 $N(\mu, \sigma^2)$ 으로부터의 확률표본이라고 할 때 확률변수 T 를

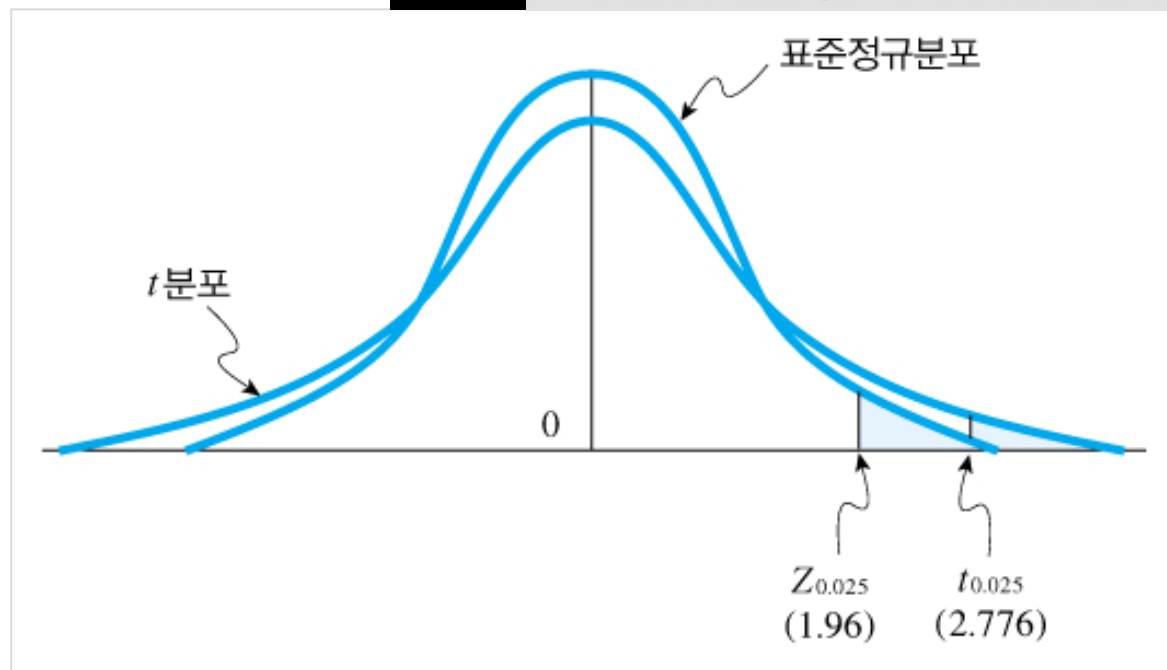
$$T = \frac{\bar{X} - \mu}{S/\sqrt{n}}$$

여기에서 $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$, $S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$ 으로 정의하면,

T 는 자유도(d.f.)가 $n-1$ 인 t 분포를 따르며 $T \sim t(n-1)$ 로 표현한다.

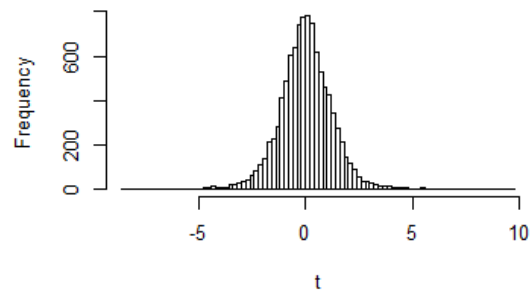


그림 표준정규분포와 $d.f. = 4$ 인 t 분포의 비교

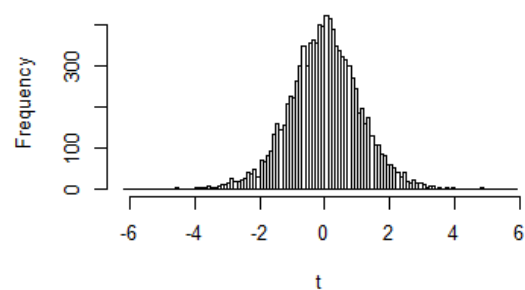


표준정규분포와 t 분포를 비교한 그림으로 t 분포가 표준정규분포보다 꼬리부분의 확률이 조금 더 큰 것을 알 수 있다.

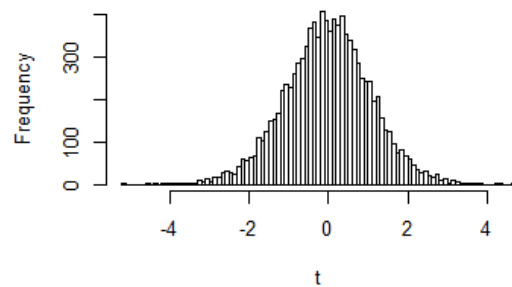
df= 5



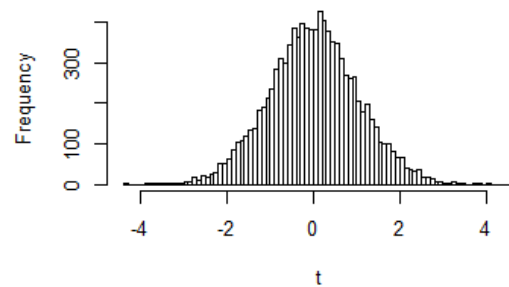
df= 10



df= 15



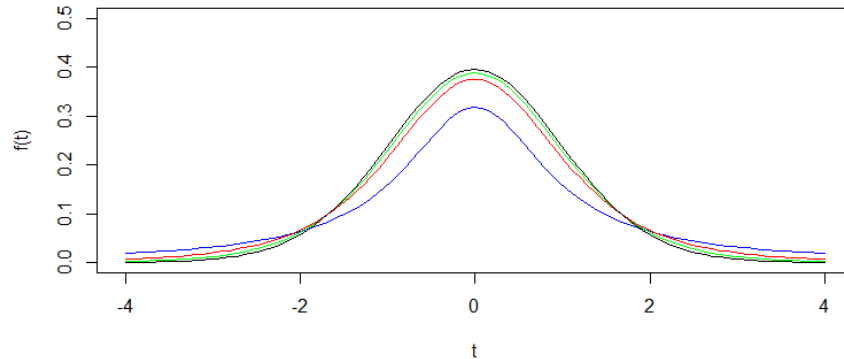
df= 30



```

1  n_list<-c(2,5,10,30) # 표본 수(n)
2  df_list<-n_list-1 # 자유도
3
4  curve(dt(x, 1, ncp=0), add=T, col="blue", xlim=c(-4, 4), ylim=c(0, 0.5), xlab="t", ylab="f(t)")
5  curve(dt(x, 4, ncp=0), add=T, col="red", xlim=c(-4, 4), ylim=c(0, 0.5), xlab="t", ylab="f(t)")
6  curve(dt(x, 9, ncp=0), add=T, col="green", xlim=c(-4, 4), ylim=c(0, 0.5), xlab="t", ylab="f(t)")
7  curve(dt(x, 29, ncp=0), add=T, col="black", xlim=c(-4, 4), ylim=c(0, 0.5), xlab="t", ylab="f(t)")
8
9  par(mfrow=c(2,2))
10
11 for (i in 1:length(df_list)) {
12
13   curve(dt(x, df_list[i], ncp=0), xlim=c(-4, 4), ylim=c(0, 0.5), xlab="t", ylab="f(t)",
14         main=paste("df=", df_list[i]))
15
16 }
17

```



t-분포는 자유도 n-1에 따라 분포의 형태가
다르므로 각각의 자유도 및 확률에 따른
t-분포의 값이 다음과 같이 주어져 있다

예 자유도가 5일 때 $t_{0.05} = 2.015$ 인데
이는 $\Pr(T \geq 2.015) = 0.05$ 을
의미한다.

```
1 t11<-rep(NA,10)
2 t12<-rep(NA,10)
3 t13<-rep(NA,10)
4 t14<-rep(NA,10)
5 t15<-rep(NA,10)
6
7 for(i in 1:10) {
8   t11[i]<-qt(0.9, i)
9 }
10 for(i in 1:10) {
11   t12[i]<-qt(0.95, i)
12 }
13 for(i in 1:10) {
14   t13[i]<-qt(0.975,i)
15 }
16 for(i in 1:10) {
17   t14[i]<-qt(0.99,i)
18 }
19 for(i in 1:10) {
20   t15[i]<-qt(0.995,i)
21 }
22
23 (tdist<-cbind(t11,t12,t13,t14,t15))
24
```

p v	0.1	0.05	0.025	0.01	0.005
1	3.078	6.314	12.706	31.821	63.657
2	1.886	2.920	4.303	6.965	9.923
3	0.1638	2.353	3.182	4.541	5.841
4	1.533	2.132	2.776	3.747	4.604
5	1.476	2.015	2.571	3.365	4.032
6	1.440	1.943	2.447	3.143	3.707
7	1.415	1.895	2.365	2.998	3.499
8	1.397	1.860	2.306	2.896	3.355
9	1.383	1.833	2.262	2.821	3.250
10	1.372	1.812	2.228	2.764	3.169
26	1.315	1.706	2.056	2.479	2.779
27	1.314	1.703	2.052	2.473	2.771
28	1.313	1.701	2.048	2.467	2.763
29	1.311	1.699	2.045	2.462	2.756
30	1.310	1.697	2.042	2.457	2.750
∞	1.282	1.645	1.960	2.326	2.576

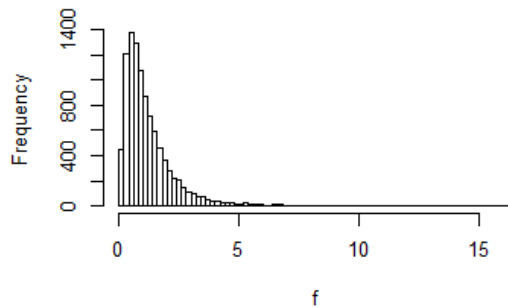
F-분포

- $X_1 = \chi^2(\nu_1), X_2 = \chi^2(\nu_2)$ 이고 X_1, X_2 이 서로 독립이며, $F = \frac{\frac{X_1}{\nu_1}}{\frac{X_2}{\nu_2}} \sim F(\nu_1, \nu_2)$

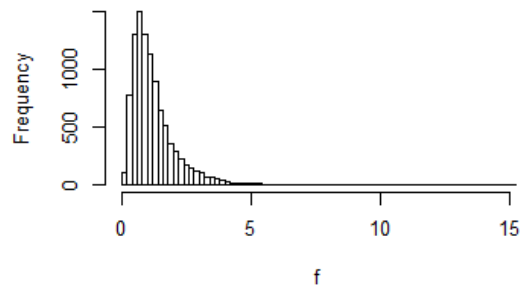
$$- F = \frac{\frac{X_{n_1-1}^2}{n_1-1}}{\frac{X_{n_2-1}^2}{n_2-1}} = \frac{\frac{(n_1-1)s_1^2}{(n_1-1)\sigma_1^2}}{\frac{(n_2-1)s_2^2}{(n_2-1)\sigma_2^2}} = \frac{\frac{s_1^2}{\sigma_1^2}}{\frac{s_2^2}{\sigma_2^2}} \sim F_{(n_1-1, n_2-1)}$$

$E(F) = \nu_2/(\nu_2-2)$ 이고, 자유도가 많아질수록 정규분포와 비슷하게 된다.

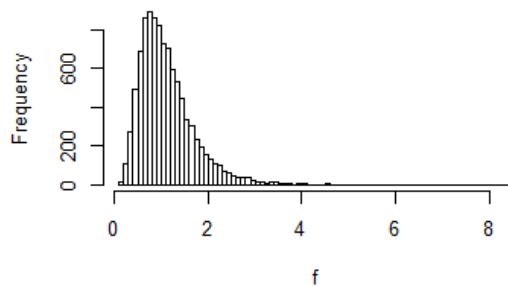
df1= 5 df2= 10



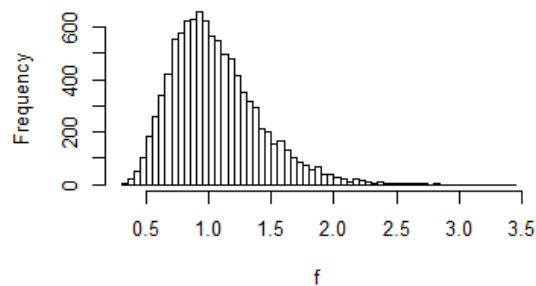
df1= 9 df2= 10



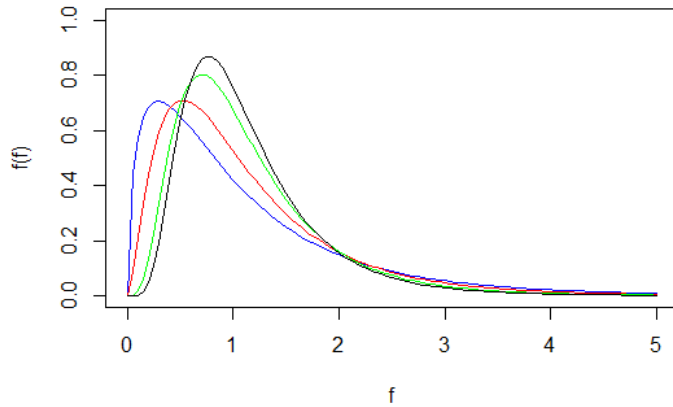
df1= 15 df2= 20



df1= 38 df2= 40



```
1 curve(df(x, 3, 15, ncp=0), add=T, col="blue", xlim=c(0,5), ylim=c(0,1), xlab="f", ylab="f(f)")
2 curve(df(x, 5, 15, ncp=0), add=T, col="red", xlim=c(0,5), ylim=c(0,1), xlab="f", ylab="f(f)")
3 curve(df(x, 10, 15, ncp=0), add=T, col="green", xlim=c(0,5), ylim=c(0,1), xlab="f", ylab="f(f)")
4 curve(df(x, 15, 15, ncp=0), add=T, col="black", xlim=c(0,5), ylim=c(0,1), xlab="f", ylab="f(f)")
5
6 df1_list<-c(3,5,10,15)
7 df2_list<-c(15,15,15,15)
8 par(mfrow=c(2,2))
9
10
11 for (i in 1:length(df1_list)) {
12   curve(df(x, df1=df1_list[i], df2=df2_list[i], ncp=0), xlim=c(0,5), ylim=c(0,1),
13         xlab="f", ylab="f(f)", main=paste("df1=", df1_list[i], "df2=", df2_list[i]))
14 }
15 |
```



F-분포는 분자의 자유도(v_1) 및 분모의 자유도(v_2)에 따라 분포의 형태가 다르므로 각각의 자유도에 따른 $\alpha = 0.05$ 에 대한 F값이 다음과 같이 주어져 있다

예 분자의 자유도가 7이고 분모의 자유도가 9인 경우 $F_{0.05} = 3.29$ 이다.

v2 v1	1	2	3	4	5	6	7	8	9	10	∞	
1	161.45	199.50	215.71	224.58	230.16	233.99	236.77	238.88	240.54	241.88	254.31	
2	18.51	19.00	19.16	19.25	19.30	19.33	19.35	19.37	19.38	19.40	19.50	
3	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.76	8.53	
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96	5.63	
5	6.61	5.79	5.41	5.19	5.05	4.95	4.48	4.82	4.77	4.74	4.37	
6	5.99	4.74	7.35	4.12	3.94	3.87	3.79	3.73	3.68	3.64	3.23	
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.64	3.23	
8	5.32	4346	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.35	2.93	
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14	2.71	
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98	2.54	
∞	3.84	3.00	2.60	2.37	2.21	2.10	2.01	1.94	1.88	1.83	1.75	1.00

```

1 f11<-rep(NA,10)
2 f12<-rep(NA,10)
3 f13<-rep(NA,10)
4 f14<-rep(NA,10)
5 f15<-rep(NA,10)
6 f16<-rep(NA,10)
7 f17<-rep(NA,10)
8
9 for(i in 1:10) {
10   f11[i]<-qf(0.95, 1, i)
11 }
12 for(i in 1:10) {
13   f12[i]<-qf(0.95, 2, i)
14 }
15 for(i in 1:10) {
16   f13[i]<-qf(0.95, 3, i)
17 }
18 for(i in 1:10) {
19   f14[i]<-qf(0.95, 4, i)
20 }
21 for(i in 1:10) {
22   f15[i]<-qf(0.95, 5, i)
23 }
24 for(i in 1:10) {
25   f16[i]<-qf(0.95, 6, i)
26 }
27 for(i in 1:10) {
28   f17[i]<-qf(0.95, 7, i)
29 }
30
31 (fdist<-cbind(f11,f12,f13,f14,f15,f16,f17))
32
> (fdist<-cbind(f11,f12,f13,f14,f15,f16,f17))
      f11      f12      f13      f14      f15      f16      f17
[1,] 161.447639 199.500000 215.707345 224.583241 230.161878 233.986000 236.768400
[2,]  18.512821  19.000000  19.164292  19.246794  19.296410  19.329534  19.353218
[3,]  10.127964   9.552094   9.276628   9.117182   9.013455   8.940645   8.886743
[4,]   7.708647   6.944272   6.591382   6.388233   6.256057   6.163132   6.094211
[5,]   6.607891   5.786135   5.409451   5.192168   5.050329   4.950288   4.875872
[6,]   5.987378   5.143253   4.757063   4.533677   4.387374   4.283866   4.206658
[7,]   5.591448   4.737414   4.346831   4.120312   3.971523   3.865969   3.787044
[8,]   5.317655   4.458970   4.066181   3.837853   3.687499   3.580580   3.500464
[9,]   5.117355   4.256495   3.862548   3.633089   3.481659   3.373754   3.292746
[10,]  4.964603   4.102821   3.708265   3.478050   3.325835   3.217175   3.135465

```