[NOTE] 단측검정 (p값은 양측검정 P값의 1/2배)

```
x1 = c ( 1.1, 2.3, 4.3, 2.2, 5.3 )        # 지혈제 A의 지혈시간
x2 = c ( 2.3, 4.3, 3.5 )                  # 지혈제 B의 지혈시간

mean(x1) ; mean(x2)

t.test(x1, x2, var.equal = T, alternative = c("less"))    #  등분산 가정하에서

T_test_2A = function(x , y) {
n1 = length(x) ; n2 = length(y)                      # 등분산 F-test
s1 = var(x) ; s2 = var(y)
F = s1 / s2
pvalue = min( 2*pf(F, n1-1, n2-1) , 2*( 1 - pf(F, n1-1, n2-1) ) )
cat( "  ============  이표본 분산비 검정  ============", "\n", "\n")
cat( "              F = " , F, " , P - value = " , pvalue , "\n", "\n")
xbar = mean(x) ; ybar = mean(y)                      # 등분산인 경우 t-test
sp = sqrt ( ( (n1 - 1) * s1 + (n2 - 1) * s2 ) / ( n1 + n2 - 2) )
T = ( xbar - ybar ) / ( sp * sqrt ( 1/n1 + 1/n2) )
pvalue = ( 1 - pt( abs(T), n1 + n2 - 2) )
cat( "  ============  이표본 평균차 검정  ============", "\n", "\n")
cat( " 등분산인 경우  : T = " , T, " , P - value = " , pvalue , "\n")

df = ( s1/n1 + s2/n2 )^2 / ( (s1/n1)^2 / (n1 - 1) + (s2/n2)^2 / (n2 - 1) )
T = ( xbar - ybar ) / sqrt ( s1/n1 + s2/n2)      # 이분산인 경우 t-test
pvalue = ( 1 - pt( abs(T), df ) )
cat( " 이분산인 경우  : T = " , T, " , P - value = " , pvalue , "\n")

}
T_test_2A(x1, x2)
```

10장. 신뢰구간에 대한 시뮬레이션

10.1. 모평균의 신뢰구간

1) 공식 : $\overline{X} \pm t_{(n-1,\alpha/2)} \frac{S}{\sqrt{n}}$

[NOTE 1] 분포값 계산(95% 기준) => P(Z < $z_{\alpha/2}$) = 0.975

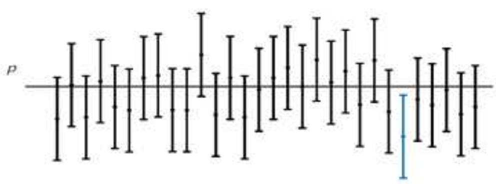
```
alpha = 0.95
a = alpha + (1 - alpha) / 2 ; a
qnorm(a)                                    # z - 분포 값
for  (i in 1 : 200) {
     ta = qt( a, i)                          # 여러 자유도에서 t - 분포 값
     cat( " df = ", i, ", ta  = " , ta, "\n")
}
```

[NOTE 2] IF 문

```
count = 0
for  (i in 1 : 10) {
     if ( i <= 5) count = count + 1
     cat( " i = ", i, ", count  = " , count, "\n")
}
count = 0
for  (i in 1 : 10) {
     if ( i >= 5 && i <= 8) count = count + 1          # && : and
     cat( " i = ", i, ", count  = " , count, "\n")
}
```

2) 비교

신뢰구간이 μ를 포함할 확률이 (1-α)$\times$100%에 가까운지 확인



3) 정규분포

- 모집단이 정규분포인 경우 모평균에 대한 신뢰구간의 신뢰도 비교 시뮬레이션

- 입력값 : 모평균(mu), 모표준편차(sd), 표본의 크기(n),
  자료 set의 크기(m), 신뢰도(alpha)

```
mu = 0 ; sd = 1 ;
n = 5 ;   m = 100000                # 난수 개수, 자료 set의 개수(넉넉히)
alpha = 0.95
```

- 프로그램

```
count = 0
a = alpha + (1 - alpha) / 2

for  (i in 1 : m) {
      x = rnorm( n, mu, sd )              # 정규분포 난수발생

      xbar = mean(x)
      se = sd(x)  / sqrt(n)

      L = xbar - qt(a, n-1) * se
      U = xbar +  qt(a, n-1) * se

        if ( mu >= L && mu <= U) count = count + 1
}

P = count / m * 100                          # % 계산

cat( "  표본의 수 = ", n, " 포함확률= " , P, "\n" )
```

4) 지수분포

- 모집단이 지수분포인 경우 모평균에 대한 신뢰구간의 신뢰도 비교 시뮬레이션
- 입력값 : 자유도(r), 표본의 크기(n), 자료 set의 크기(m), 신뢰도(alpha)

```
r = 1
n = 5 ;  m = 100000        # 난수 개수, 자료 set의 개수(넉넉히)
alpha = 0.95
```

- 프로그램

```
mu = 1/r
count = 0
a = alpha + (1 - alpha) / 2

for  (i in 1 : m) {
        x = rexp(n, rate = r)                    # 지수분포 난수발생

        xbar = mean(x)
        se = sd(x)  / sqrt(n)

        L = xbar - qt(a, n-1) * se
        U = xbar +  qt(a, n-1) * se

          if ( mu >= L && mu <= U) count = count + 1
}

P = count / m * 100                              # % 계산

cat( "  표본의 수 = ", n, " 포함확률= " , P, "\n" )
```

5) 포아송분포

● 모수

```
lambda = 5
n = 5 ;   m = 100000          # 난수 개수, 자료 set의 개수(넉넉히)
alpha = 0.95
```

● 프로그램

```
mu = lambda
count = 0
a = alpha + (1 - alpha) / 2

for  (i in 1 : m) {
        x = rpois(n, lambda)                       # 포아송 분포 난수발생

        xbar = mean(x)
        se = sd(x)  / sqrt(n)

        L = xbar - qt(a, n-1) * se
        U = xbar +  qt(a, n-1) * se

          if ( mu >= L && mu <= U) count = count + 1
}

P = count / m * 100                                # % 계산

cat( "  표본의 수 = ", n, " 포함확률= " , P, "\n" )
```

[과제27] 모수를 다르게

1. 정규분포
2. 지수분포
3. 포아송분포

1) m : 충분히(1000000), n : 다양하게 (5, 20, 50) 모의실험하고

2) 분포에 따라 표본의 크기에 따라 95%로 얼마나 빨리 다가가는지 비교 및 해석