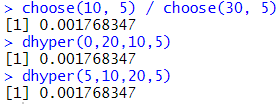
**[컴퓨터과학과 2018301034 서 용 주]**

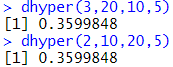
**4장 연습문제**

**15번.** 우리 회사의 프린트는 A, B 2대가 있다. 서버 A에는 20명이 연결되어 있고, 서버 B에는 10명이 연결되어 있다. 연결된 직원 5명을 추출하였을 때 아래 질문들을 R로 구현하여라.

**(1) 서버 A에 연결된 직원이 한명도 뽑히지 않을 확률은?**



**(2) 서버 A에 연결된 직원이 3명 뽑힐 확률은?**



**(3) 서버 A에 연결된 직원이 3명 이상 뽑힐 확률은?**

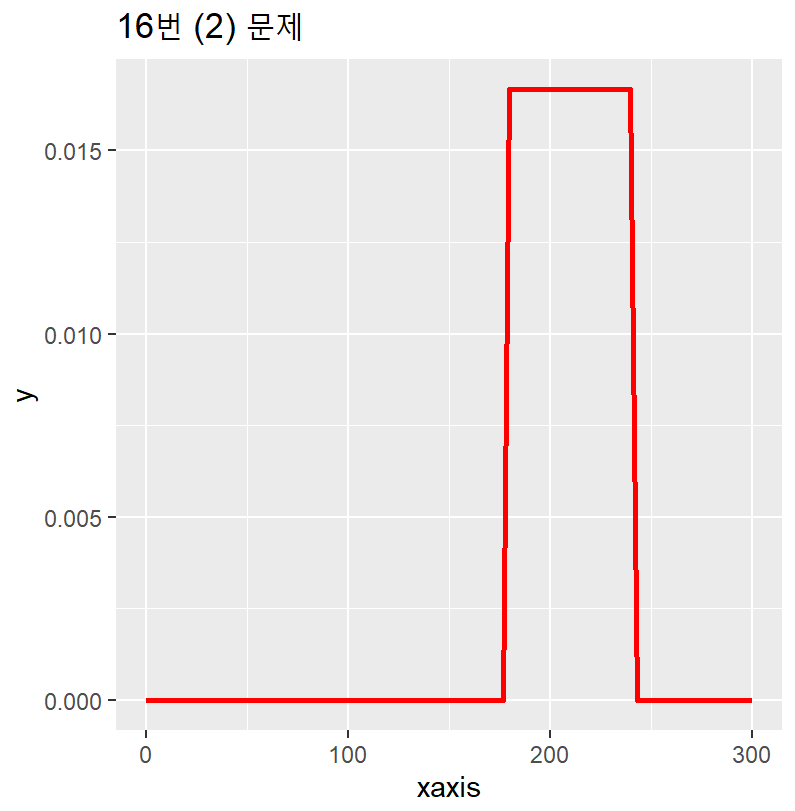
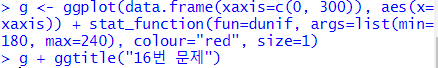


**16번.** 우리 회사의 인터넷 이용자들의 이용시간은 3분~4분 사이에 랜덤하게 이용한다. 아래 물음들을 R로 구현하여라.

**(1) 확률밀도함수는?**



(**2) 확률밀도 함수 그래프를 그려라**



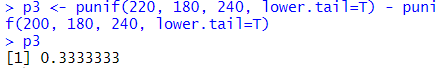
**(3) 인터넷 이용시간이 3분 10초 이하일 확률은?**



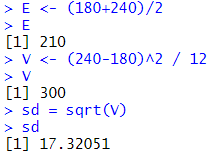
**(4) 인터넷 이용시간이 3분 30초 이상일 확률은?**



**(5) 인터넷 이용시간이 3분 20초에서 3분 40초 사이일 확률은?**



**(6) 평균과 분산, 표준편차를 구하여라.**



**18번.** 변수z는 표준정규 확률변수이다. 다음의 확률들을 R로 계산하여라.

**(1) P(z >-1.53)**



**(2) P(z < 1.53)**



**(3) P(-2.1 < z < 2.11)**

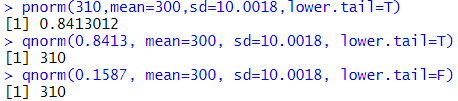
**(4) P(-1.96 < z < 1.96)**



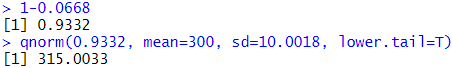
**20번.** 수능점수의 평균은 300점이고, 수능점수의 분포는 정규분포라고 하자. 아래 물음들을 R로 계산하여라.

**(1) sd = 10.0018**

**(2)**



**(3)**

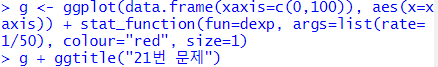
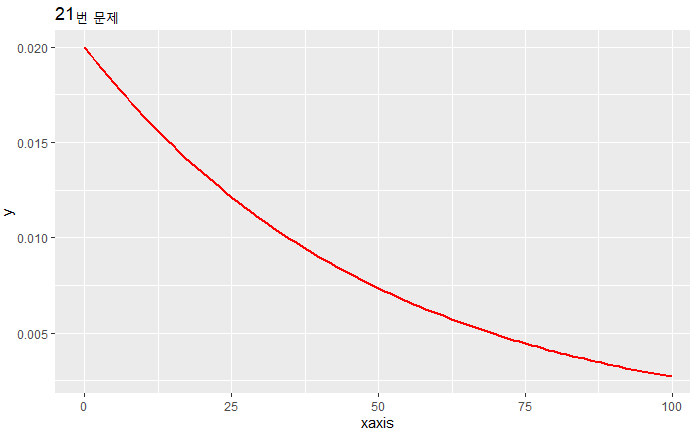


**21번.** 우리 회사 네트워크 프린터에 다음 요청이 들어오기 까지 평균 50초가 소요된다. 아래 물음들을 R로 구현하여라

**(1)**



**(2)**



**(3)**



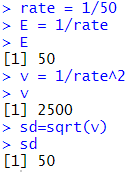
**(4)**



**(5)**



**(6)**



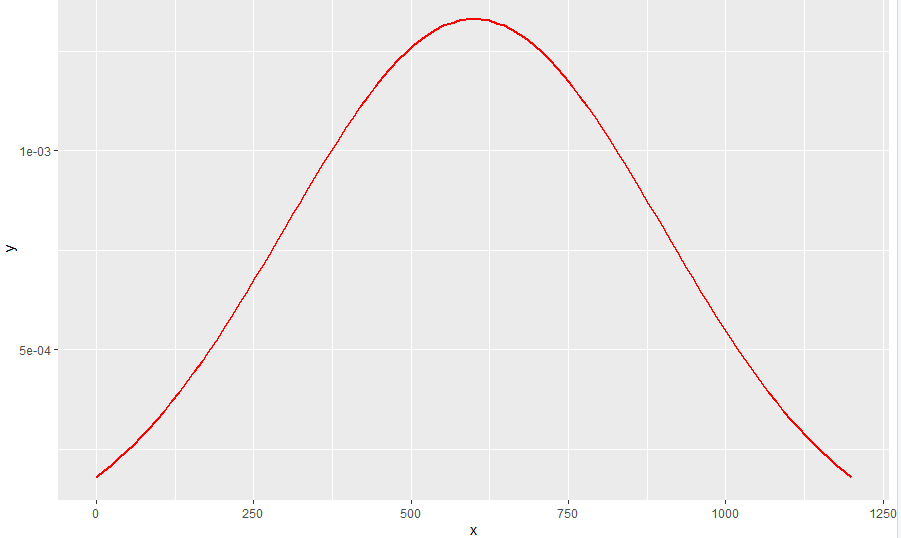
**[4장 실습과제]**

우리 회사의 동영상 가입자는 100만여명이다. 동영상 이용자들을 분석해 보니 평균 10분, 표준편차 5분을 이용하고 있다. 아래 물음들을 R로 구현하여라.

**(1)**



**(2)**



**(3)**



**(4)**



**(5)**



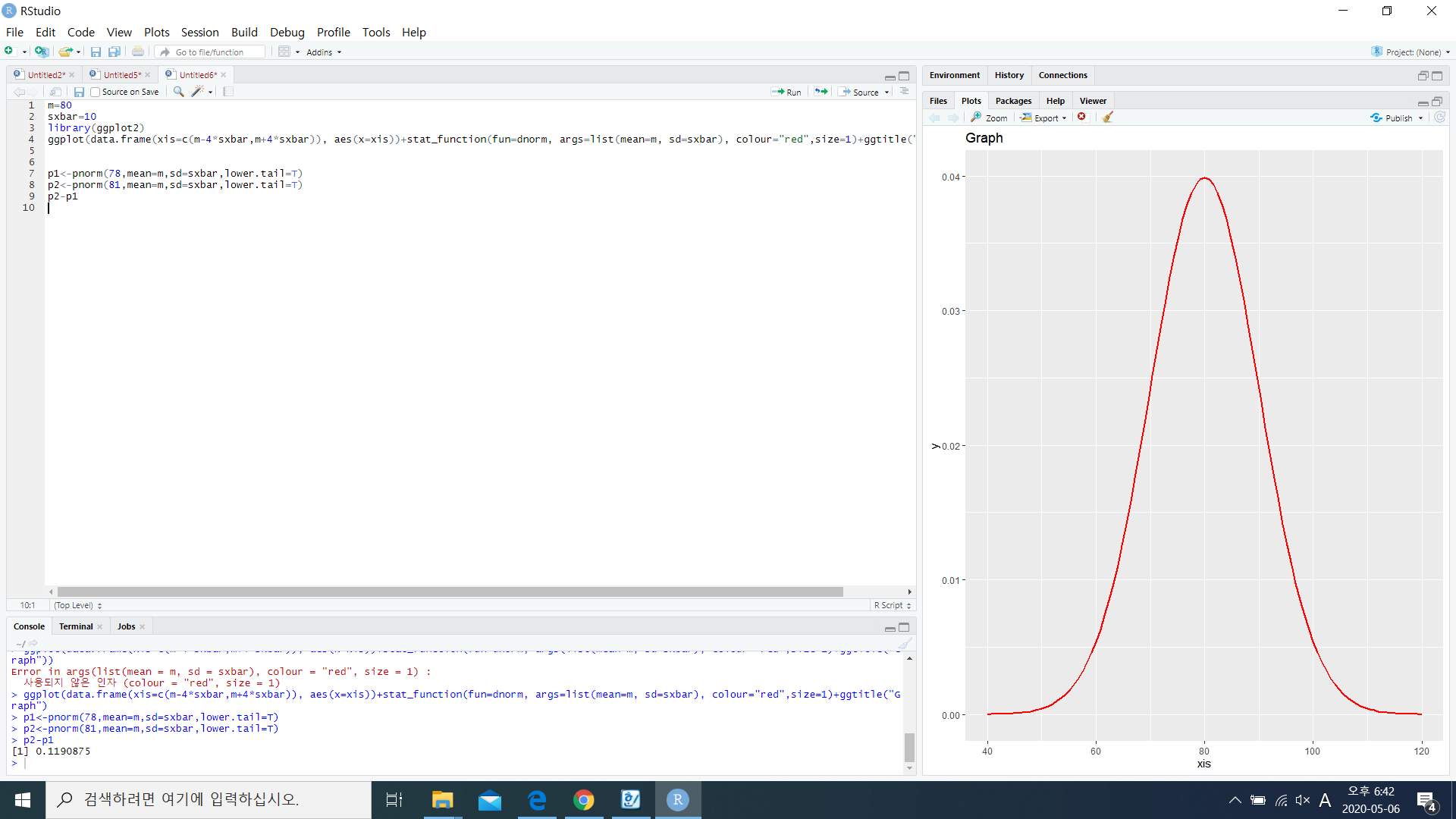
**(6)평균 = 10분, 분산 = 25분, 표준편차 = 5분**

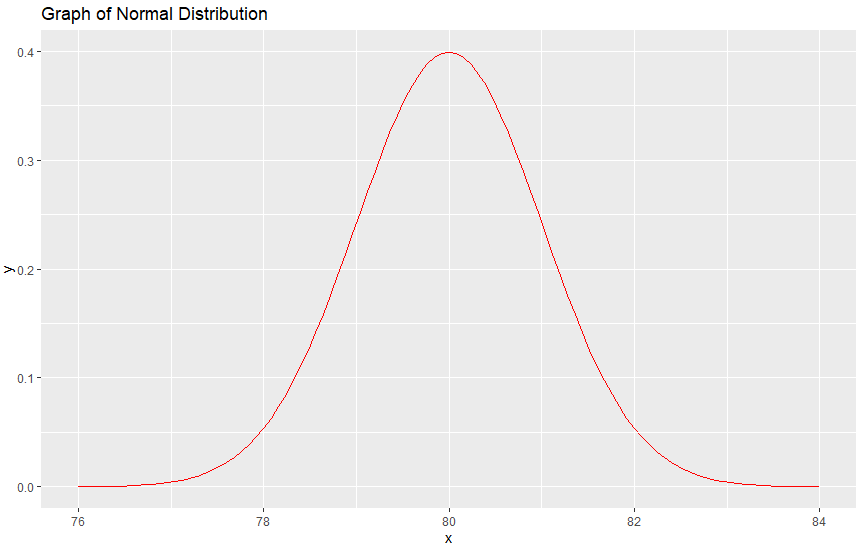
**5장 연습문제**

**4번.** 우리 회사 직원의 자바시험 성적은 평균이 80이고, 표준편차가 10인 정규분포를 한다. 이를 확인하기 위해 100명의 직원을 뽑아서 구한 평균이 78보다 크고 81보다 적을 확률은?

또한 확률밀도 함수를 구하고, 이에 대한 그래프를 그려라.

**-> 평균이 78보다 크고 81보다 적을 확률.**

 **-> 이를 R 이용하여 그래프로 나타내면 다음과 같다.**

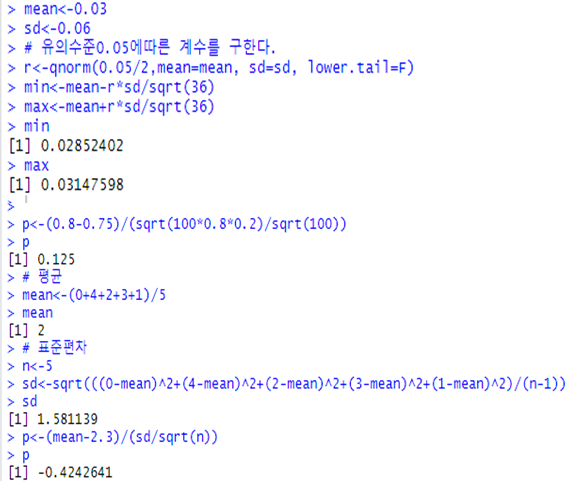


**6번.** 회사 서버성능을 측정하기 위해 36번 응답시간을 측정해보니, 평균 0.03초이었고 표준편차가 0.06초이었다. 우리 회사 서버의 응답시간에 대한 95% 신뢰구간을 구하여라.

**-> 표준편차를 아는 경우이다. 표본평균을 Xn이라고 잡으면, Xn은 모평균 를 평균으로 하고, 표준편차의 경우에는 무한모집단이므로, 0.06/6 으로 하는 정규분포를 따른다.**

**이때, 모평균 의 범위는 (Xn-Z\*0.01, Xn+Z\*0.01)인데, 95% 신뢰구간이면 표준정규분포에서 z=1.96인 경우이므로, Z값을 1.96으로 놓는다.**

**Xn과 Z를 각각 대입하면 (0.03-1.96\*001, 0.03+1.96\*0.01)이 된다.**



**7번.** 어느 투표에서 후보자 A와 B가 나왔다. 유권자는 10만여 명인데, 지지율을 확인하기 위해 여론조사를 실시하였따. 1000명을 조사하여보니 후보자 A에 대한 지지율이 45%였다. 후보자 A와 B의 지지율에 대한 99% 신뢰구간을 구하여라. 후보자 A가 후보자 B에 열위에 있는가. 이에 대해 논하여라.

**표본비율에 대한 중심극한정리에 따라 np>=5이고, nq>=5이므로 정규분포에 근사된다.**

**또한 유한모집단이지만, N=10만, n=1000으로 n/N <0.05이므로 표준오차 값으로 root(pq/n)을 사용해도 좋다. 그러면 A의 지지율을 Pa라 할 때 Pa는 m= 0.45이고, sd= root(0.45\*0.55/1000)인 정규분포이므로, 신뢰도 99% 구간은 (0.45 - 2.58\*sqrt(0.45\*0.55/1000), 0.45 + 2.58\*sqrt(0.45\*0.55/1000) 이다.**

**B의 지지율 Pb는 마찬가지로 구하면, 단순히 m=0.55라고 하면 되므로 그 구간은**

**(0.55 - 2.58\*sqrt(0.45\*0.55/1000), 0.55 + 2.58\*sqrt(0.45\*0.55/1000)이다.**

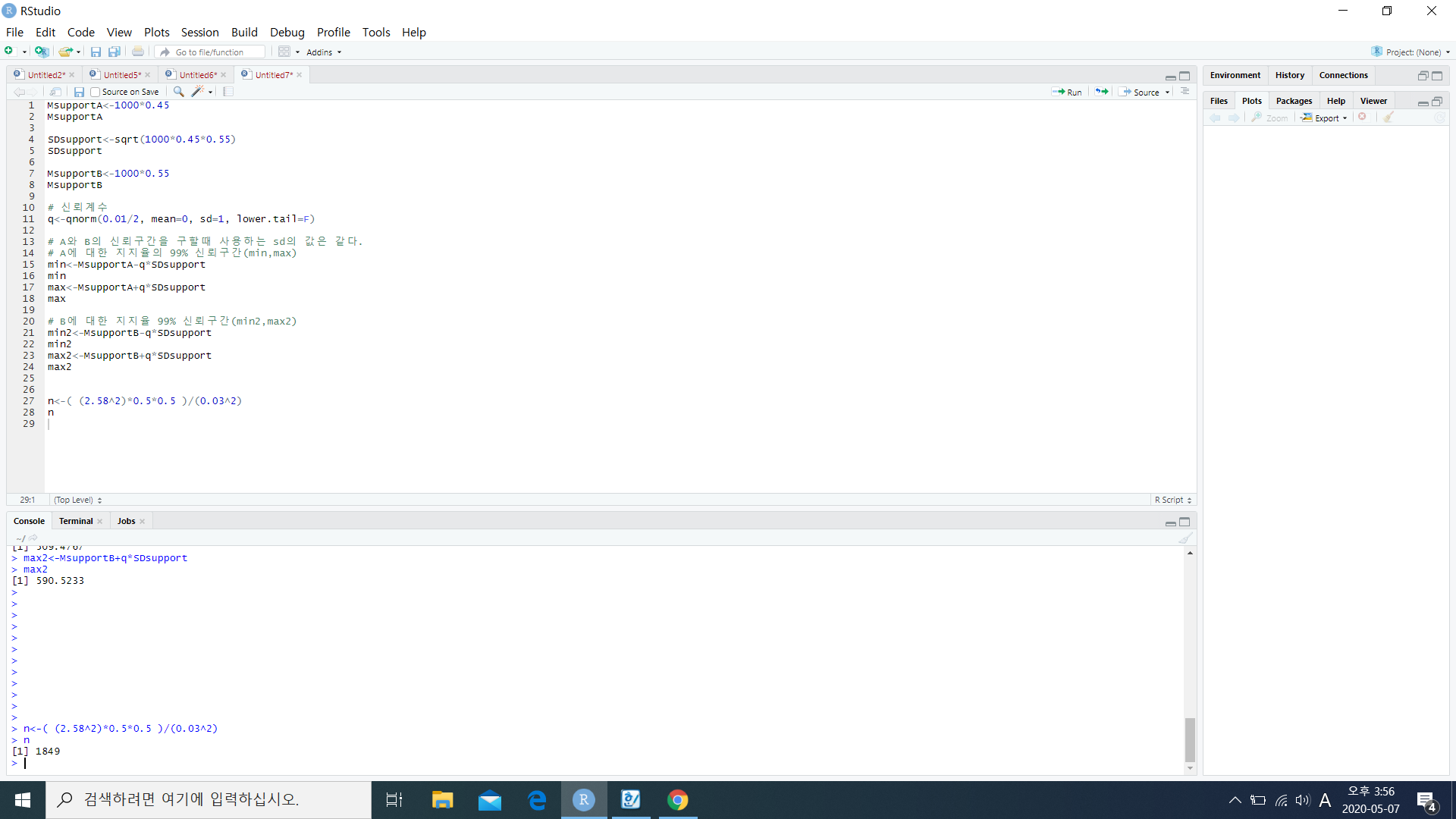
**(0.4094 <= a <= 0.4905)**

**(0.5094 <= b <= 0.5905)**

**a의 최대 지지율인 49%가 b의 최소 지지율인 50.94%보다 낮기 때문에 확실히 열세라고 볼 수 있다.**

**8번.** 어느 투표에서 지지율에 대한 조사를 실시하려고 한다. 99%의 신뢰구간을 3% 이내로 관리하기 위해서는 몇 명의 표본을 추출하여야 할까?

**-> 1849명**



**11번.** 우리는 서버의 처리능력이 향상된 소프트웨어를 개발하였다. 기존의 서버는 초당 평균 10000건을 처리하였다. 그런데 우리의 소프트웨어를 탑재하여 36회 실험하니 평균 10005건을 처리하였다. 표준편차는 10회이었다. 우리의 소프트웨어가 처리능력을 향상시켰다고 할 수 있는가? 유의수준 0.05로 p값을 이용하여 5단계 가설검정을 실시하여라.

**1) 가설설정**

**귀무가설 H0 : μ = 10000**

**대립가설 H1 : μ > 10000**

**2) 유의수준 α = 0.05**

**3) 검정통계량**

 **4) p값 계산**

** 5) p값 = 0.002474416 < 0.05 이다. 따라서 유의수준 5%하에서 처리 능력이 향상되었다 볼 수 있습니다. (귀무가설을 기각)**

**12번.** 우리 회사는 새로운 검색엔진을 개발하였다. 기존의 엔진은 컨텐츠를 검색하는 데 1초 걸렸다. 새로운 엔진으로 25회 검색해 보니 평균 0.6초 표준편차는 1초 걸렸다. 새로운 엔진이 빠르다고 할 수 있는가? 유의수준 5%로 p값을 이용하여 5단계 가설검정을 실시하여라.

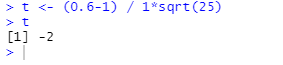
**1) 가설설정**

**귀무가설 H0 : μ = 1**

**대립가설 H1 : μ < 1**

**2) 유의수준 α = 0.05**

**3) 검정통계량**

**4) p값 계산**

**5) p값 = 0.9715301 > 0.05 이다. 따라서 유의수준 5%하에서 새로운 엔진이 빠르다고 볼 수 없다. (귀무가설을 채택)**

**[5장 실습과제]**

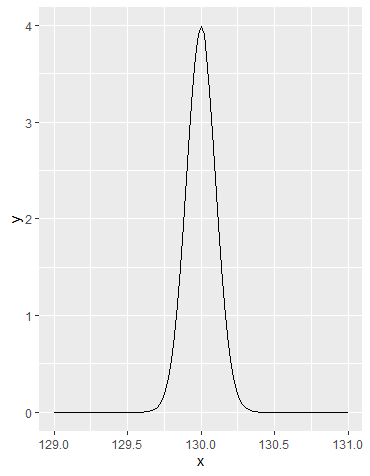
어느 인터넷 서비스 회사에서 이용자는 평균 2분이라고 알려져 있다. 이를 확인하기 위해 100명을 조사해보니 평균 2분 10초, 표본표준편차는 0.1초였다. 아래 물음에 답하여라.

**(1) 표본분포의 확률밀도함수는?**

f(x) = 1 / 0.1\*sqrt(2π) \* e ^ (1/2 \* ( (x-130)^2 / 0.1^2) ) )

**(2) 확률밀도 함수 그래프를 그려라.**

**ggplot(data.frame(x=c(129, 131)), aes(x=x)) + stat\_function(fun=dnorm, args=list(mean=130, sd=0.1))**



**(3)**

**1) 가설설정**

**귀무가설 H0 : μ = 120**

**대립가설 H1 : μ =/ 120**

**2) 유의수준 α = 0.05**

**3) 검정통계량**

**4) p값 계산**

**5) p값 = 40840172e^-200 < 0.05 이다. 따라서 유의수준 5%하에서 인터넷 서비스 이용자는 평균 2분이라고 할 수 없다. (귀무가설 기각)**