1. $X \sim B(100, 0.55)$일 때, 확률 $P[X \le 45]$을 다음 방법들을 통하여 구하여라.
   (1) 연속성 수정을 사용하지 않은 정규 근삿값
   (2) 연속성 수정을 사용한 정규 근삿값

(sol)

$np = 100 \times 0.55 = 55$, $npq = 100 \times 0.55 \times 0.45 = 24.75$이므로

$$X \approx N(55, (4.97)^2)$$

이다.

(1) $P(X \le 45) = P\left(\frac{X-55}{4.97} \le \frac{45-55}{4.97}\right) \approx P(Z \le -2.01) = 0.0222$

(2) $P(X \le 45) = P(X \le 45.5) = P\left(\frac{X-55}{4.97} \le \frac{45.5-55}{4.97}\right) \approx P(Z \le -1.91) = 0.281$

※ $P(X \le 45) = \sum_{x=0}^{45} \binom{100}{x} (0.55)^x (0.45)^{100-x} = 0.0284$

2. (1) 확률변수 X와 Y가 서로 독립이고 자유도가 각각 3과 6인 카이제곱분포를 따를 때, $P(X+Y \ge 12)$일 확률을 구하여라.
   (2) 두 분포 $X \sim N(0,1)$과 $Y \sim t(n)$에 대하여,
   확률 $P(X \ge 1)$와 $P(Y \ge 1)$를 다음에 주어진 t분포의 자유도 n의 크기에 따라 그 값을 비교하여라.
   ① $n=5$ ② $n=15$ ③ $n=30$ ④ $n=100$
   (3) F분포 $F \sim F(n_1, n_2)$에 대하여, 다음을 구하여라.
   ① $n_1=7, n_2=15$일 때, $f_{0.95}(n_1, n_2)$ ② $n_1=24, n_2=9$일 때, $f_{0.025}(n_1, n_2)$ ③ $n_1=13, n_2=9$일 때, $P(F \ge 12)$

(sol)

(1) $X \sim \chi^2(3)$, $Y \sim \chi^2(6)$이고, X와 Y가 서로 독립이므로 $X+Y \sim \chi^2(9)$이다.

$P(X+Y \ge 12) = 0.2133$

※ $P(X+Y \ge 11.39) = 0.25$, $P(X+Y \ge 12.24) = 0.20$이므로,

$$\frac{0.2-0.25}{12.24-11.39} \approx \frac{0.2-\alpha}{12.24-12}$$

$\Rightarrow$ $0.85(0.2-\alpha) \approx 0.012$

$\Rightarrow$ $0.2-\alpha \approx 0.0141$

$(\therefore)$ $\alpha \approx 0.2141$

(2) $P(X \ge 1) = 1 - P(X < 1) = 1 - 0.8413 = 0.1587$

① $t_\alpha(5) = 1 \Rightarrow \alpha = 0.1816$

② $t_\alpha(15) = 1 \Rightarrow \alpha = 0.1666$

③ $t_\alpha(30) = 1 \Rightarrow \alpha = 0.1627$

④ $t_\alpha(100) = 1 \Rightarrow \alpha = 0.1599$

(3) ① $f_{0.95}(7,15) = \frac{1}{f_{0.05}(15,7)} = \frac{1}{3.51} = 0.2849$

② $f_{0.025}(24,9) = \frac{1}{f_{0.05}(15,7)} = \frac{1}{3.51} = 3.6142$

③ $f_\alpha(13,9) = 12 \Rightarrow \alpha = 0.00039$

3. 확률분포 $f_X(1)=0.6$, $f_X(2)=0.4$를 잡는 모집단으로부터 크기 4인 표본을 임의 추출하였을 때, 표본평균 $\overline{X}$의 확률분포와 평균 그리고 분산을 구하여라.

(sol)

$\overline{X}=\frac{4}{4}$: $\{(1,1,1,1)\}$

$\overline{X}=\frac{5}{4}$: $\{(1,1,1,2),(1,1,2,1),(1,2,1,1),(2,1,1,1)\}$

$\overline{X}=\frac{6}{4}$: $\{(1,1,2,2),(1,2,1,2),(2,1,1,2)(1,2,2,1),(2,1,2,1),(2,2,1,1)\}$

$\overline{X}=\frac{7}{4}$: $\{(1,2,2,2),(2,1,2,2),(2,2,1,2),(2,2,2,1)\}$

$\overline{X}=\frac{8}{4}$: $\{(2,2,2,2)\}$

$\overline{X}$	$\frac{4}{4}$	$\frac{5}{4}$	$\frac{6}{4}$	$\frac{7}{4}$	$\frac{8}{4}$
$f_{\overline{X}}$	$\binom{4}{0}(0.6)^4(0.6)^0=0.1296$	$\binom{4}{1}(0.4)^3(0.6)^1=0.3456$	$\binom{4}{2}(0.4)^2(0.6)^2=0.3456$	$\binom{4}{3}(0.4)^1(0.6)^3=0.1536$	$\binom{4}{4}(0.4)^0(0.6)^4=0.0256$

$E[X]=\sum_{i=1}^{2} x_i f_X(x_i)=1(0.6)+2(0.4)=1.4$,

$E[X^2]=\sum_{i=1}^{2} {x_i}^2 f_X(x_i)=1(0.6)+4(0.4)=2.2$

$Var[X]=E[X^2]-(E[X])^2=2.2-(1.4)^2=0.24$

$(\therefore)$ $E[\overline{X}]=E[\frac{1}{4}\sum_{i=1}^{4} X_i]=\frac{1}{4}\sum_{i=1}^{4} E[X_i]=1.4$,

$Var[\overline{X}]=Var[\frac{1}{4}\sum_{i=1}^{4} X_i]=\frac{1}{16}\sum_{i=1}^{4} Var[X_i]=0.06$

※ $E[\overline{X}]=\sum_{i=1}^{5} \overline{x}_i f_{\overline{X}}(\overline{x}_i)=\frac{4}{4}\times 0.1296+\frac{5}{4}\times 0.3456+\frac{6}{4}\times 0.3456+\frac{7}{4}\times 0.1536+\frac{8}{4}\times 0.0256=1.4$

$Var[\overline{X}]=\sum_{i=1}^{5}(\overline{x}_i-\overline{x})^2 f_{\overline{X}}(\overline{x}_i)$

$=(\frac{4}{4}-1.4)^2\times 0.1296+(\frac{5}{4}-1.4)^2\times 0.3456+(\frac{6}{4}-1.4)^2\times 0.3456+(\frac{7}{4}-1.4)^2\times 0.1536+(\frac{8}{4}-1.4)^2\times 0.0256=0.06$

4. 대도시의 분주한 교차로를 통과하는 자동차들이 신호등 앞에서 대기하는 시간을 조사하기 위하여 30개의 교차로를 임의로 선정하여 조사한 결과 다음 표와 같은 결과를 얻었다. 이때 표본평균과 표본분산과 표본표준편차를 구하여라.

0.8	3.3	1.2	1.3	2.4	2.2	2.0	2.1	3.1	1.2	3.0	2.3	3.1	5.3	3.4
2.5	3.1	2.4	2.8	3.0	1.9	3.7	3.7	2.4	1.5	3.1	2.6	3.7	3.8	2.4

(단위 : 분)

(sol)

$\overline{x}=\frac{1}{30}\sum_{i=1}^{30} x_i=\frac{1}{30}\times 79.3=2.6433$,

$s^2=\frac{1}{29}\sum_{i=1}^{30}(x_i-\overline{x})^2=\frac{1}{29}\times 26.3737=0.9094 \Rightarrow s=\sqrt{0.9094}=0.9536$

5. 하루 동안에 스톡옵션의 가격이 1만원 오를 확률이 0.52, 1만원 내릴 확률이 0.48을 갖는다고 하자.

첫 날 200만원을 투자하여 100일 후의 가격이 $X=200+\sum_{i=1}^{100} X_i$ 일 때, 다음을 구하여라.

(1) $E[X_i]$와 $Var[X_i]$를 구하여라.

(2) 중심극한 정리에 의하여 100일 후의 가격이 210만원 이상일 확률을 구하여라.

(sol)

(1) $E[X_i]=1(0.52)+(-1)(0.48)=0.04,\ E[X^2]=1(0.52)+(-1)^2(0.48)=1$

$Var[X_i]=E[X_i{}^2]-(E[X_i])^2=1-(0.04)^2=0.9984$

$(\therefore)\ E[X]=E[200+\sum_{i=1}^{100} X_i]=200+\sum_{i=1}^{100} E[X_i]=204$

$Var[X]=Var[200+\sum_{i=1}^{100} X_i]=\sum_{i=1}^{100} Var[X_i]=99.84$

(2) $\overline{X}=\frac{1}{100}\sum_{i=1}^{100} X_i$이라 하면, 중심극한정리에 의하여 $\overline{X}\approx N(0.04, 0.9984)$이다.

따라서 $Y=100\overline{X}=\sum_{i=1}^{100} X_i$이면, $Y\approx N(4, 99.84)$이다.

$P(200+\sum_{i=1}^{100} X_i \geq 210)=P(Y=\sum_{i=1}^{100} X_i \geq 10)=P(Z\geq \frac{10-4}{\sqrt{99.84}})=1-P(Z\leq 0.6)=1-0.7257=0.2743$

6. 정상인의 경우 혈청 속에 포함된 철분 함량은 혈액 100ml당 평균과 표준편차가 각각 120ml, 15ml인 정규분포를 따른다고 한다. 정상인 25명을 임의로 선발하여 검사한 철분함량의 평균이 95ml이상 105ml이하일 확률을 구하여라.

(sol)

$X\sim N(120, 15^2)$이므로,

표본크기 25인 표본평균 $\overline{X}$는 중심극한정리에 의하여 평균이 120이고, 표준편차 $\frac{15}{\sqrt{25}}=3$인 정규분포를 따른다.

즉, $\overline{X}\sim N(120, 3^2)$이다.

$(\therefore)\ P(95\leq \overline{X}\leq 105)=P(\frac{95-120}{3}\leq Z\leq \frac{105-120}{3})=P(-\frac{25}{3}\leq Z\leq -5)=0.00000029$

7. 모평균이 $\mu_1=550$, $\mu_2=500$이고 모표준편차가 $\sigma_1=9$, $\sigma_2=16$인 두 정규모집단에서 각각 크기 50과 40인 표본을 임의로 추출하였을 때, 두 표본평균의 차가 48이상 52이하일 확률을 구하여라.

(sol)

$X_1\sim N(550, 9^2)$, $X_2\sim N(500, 16^2)$이므로, $\overline{X}_1\sim N(550, \frac{9^2}{50})$, $X_2\sim N(500, \frac{16^2}{40})$이다.

따라서 $\overline{X}_1-\overline{X}_2\sim N(50, \frac{9^2}{50}+\frac{16^2}{40}=2.83^2)$이다.

$(\therefore)(\therefore)\ P(48\leq \overline{X}_1-\overline{X}_2\leq 52)=P(\frac{48-50}{2.83}\leq Z\leq \frac{52-50}{2.83})=P(|Z|\leq 0.71)=2(0.7611-0.5)=0.5222$

8. 모평균 μ인 정규모집단으로부터 크기 25인 표본을 임의추출할 때, $P(\frac{|\overline{X}-\mu|}{S}<k)=0.95$를 만족하는 상수 k를 구하여라.

(sol)

$\frac{\overline{X}-\mu}{\frac{S}{\sqrt{n}}}\sim t(n-1)$이므로, $\frac{\overline{X}-\mu}{\frac{S}{\sqrt{25}}}\sim t(24)$

$P(\frac{|\overline{X}-\mu|}{S}<k)=P(5\frac{|\overline{X}-\mu|}{S}<5k=2.064)=0.95 \Rightarrow k=0.4128$

※ $P(T\geq 2.064)=0.025$

따라서 $\overline{X}_1-\overline{X}_2\sim N(50, \frac{9^2}{50}+\frac{16^2}{40}=2.83^2)$이다.

$(\therefore)\ P(48\leq \overline{X}_1-\overline{X}_2\leq 52)=P(\frac{48-50}{2.83}\leq Z\leq \frac{52-50}{2.83})=P(|Z|\leq 0.71)=2(0.7611-0.5)=0.5222$

9. 두 정규모집단 A와 B의 모분산은 동일하고, 평균은 각각 $\mu_X = 700$, $\mu_Y = 680$이라 한다. 이 때 두 모집단으로부터 표본을 추출하여 다음과 같은 결과를 얻었다. 이 때 단위는 mg이다.

A표본 : $n_1 = 17$, $\overline{x} = 704$, $s_X = 39.25$

B표본 : $n_2 = 10$, $\overline{y} = 675$, $s_Y = 43.75$

(1) 합동표본분산의 관측값 ${s_p}^2$를 구하여라.

(2) 두 표본평균의 차 $T = \overline{X} - \overline{Y}$에 대한 확률분포를 구하여라.

(3) $P(T \geq t_0) = 0.05$인 t_0를 구하여라.

(sol)

(1) ${s_p}^2 = \dfrac{1}{n_1 + n_2 - 2}\{(n_1 - 1){s_X}^2 + (n_2 - 1){s_Y}^2\} = \dfrac{1}{25}\{16(39.25)^2 + 9(43.75)^2\} = 1675.0255 \Rightarrow s_p = 40.927$

(2) $\dfrac{\overline{X} - \overline{Y} - (\mu_1 - \mu_2)}{S_p\sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}} \sim t(n_1 + n_2 - 2)$이므로, $\dfrac{T - 20}{(40.927)\sqrt{\dfrac{27}{170}}} = \dfrac{T - 20}{16.31} \sim t(25)$이다.

(3) $P(T \geq t_0) = P(\dfrac{T - 20}{16.31} \geq \dfrac{t_0 - 20}{16.31} = 1.708) = 0.05 \Rightarrow t_0 = 20 + 27.86 = 47.86$

10. 어느 회사에서 생산된 배터리의 10%가 불량품이라고 한다. 이 회사에서 임의로 10개를 선정하였을 때, 다음을 구하여라.

(1) 불량품이 없을 확률

(2) 2개 이상 불량품이 포함될 확률

(3) 15% 이상 불량품이 포함될 확률

(sol)

(1) 불량품의 수를 X라 하면 $X \sim B(10.0.1)$이다.

$P(X = 0) = \binom{10}{0}(0.9)^{10}(0.1)^0 = 0.3487$

(2) $P(X \geq 2) = 1 - P(X \leq 1) = 1 - 0.7361 = 0.2639$

$P(X = 0) + P(X = 1) = \binom{10}{0}(0.9)^{10}(0.1)^0 + \binom{10}{1}(0.9)^9(0.1)^1 = 0.3487 + 0.3874 = 0.7361$

(3) $P(X \geq 1.5) = P(X \geq 2) = 0.2639$

11. 트랜스미션에 대한 주요 결점은 외부에 의한 영향으로 발생하며, 과거의 경험에 의하면 모든 결점의 75%가 번개에 의한 것으로 알려졌다. 이를 확인하기 위하여 300개의 결함이 있는 트랜스미션을 조사한 결과 200개 이상 번개에 의한 원인일 확률을 구하여라.

(sol)

$\hat{p} \approx N(0.75, \dfrac{(0.75)(0.25)}{300} = 0.025^2)$

$P(\hat{p} \geq \dfrac{200}{300}) = P(Z \geq \dfrac{\frac{2}{3} - 0.75}{0.025} = -3.33) = 1 - P(Z \leq 3.33) = 1 - 0.9996 = 0.0004$