6. Perform the following using uniform distribution between 200 and 240

A. P(x > 230)

B. P(205 ≤ x ≤ 220)

```
lower_bound <- 200
upper_bound <- 240
x1 <- 230
probability_x1 <- 1 - punif(x1, lower_bound, upper_bound)
x2_min <- 205
x2_max <- 220
probability_x2 <- punif(x2_max, lower_bound, upper_bound) - punif(x2_min, lower_bound,
upper_bound)
cat("a. P(x > 230):", probability_x1, "\n")
cat("b. P(205 ≤ x ≤ 220):", probability_x2, "\n")
```

Output:

a. P(x > 230): 0.25

b. P(205 ≤ x ≤ 220): 0.375

Manual:

To calculate probability using uniform distribution between 200 and 240.
We can calculate using this formula

$$P(x) = \frac{maximum\ value - minimum\ value}{upper\ limit - lower\ limit}$$

a) P(x>230)
Maximum value is not given, so maximum value is equal to upper limit

maximum value = 240
minimum value = 230
upper limit = 240
lower limit = 200

$$P(x > 230) = \frac{240 - 230}{240 - 200} = 0.25$$

The probability that x is greater than 230 in this uniform distribution is 0.25 or 25%

c) $P(205 \leq x \leq 220)$
maximum value = 220

minimum value = 205
upper limit = 240
lower limit = 200

$$P(205 <= x <= 220) = \frac{220 - 205}{240 - 200} = 0.375$$

The probability that x is between 205 and 220 is 0.375 or 37.5%

7. Following are the scores of max vertical jumps before and after the training program. Test whether the training program is helpful to the students (Use Paired t-test).

Player	Max Vertical Jump Before Training Program	Max Vertical Jump After Training Program
Player 1	22	24
Player 2	20	22
Player 3	19	19
Player 4	24	22
Player 5	25	28
Player 6	25	26
Player 7	28	28
Player 8	22	24
Player 9	30	30
Player 10	27	29
Player 11	24	25
Player 12	18	20
Player 13	16	17
Player 14	19	18

Player 15	19	18
Player 16	28	28
Player 17	24	26
Player 18	25	27
Player 19	25	27
Player 20	23	24

```
before_training <- c(22, 20, 19, 24, 25, 25, 28, 22, 30, 27, 24, 18, 16, 19, 19, 28, 24, 25, 25, 23)

after_training <- c(24, 22, 19, 22, 28, 26, 28, 24, 30, 29, 25, 20, 17, 18, 18, 28, 26, 27, 27, 24)

result <- t.test(before_training, after_training, paired = TRUE)

print(result)

if (result$p.value < 0.05)

{

cat("The training program is statistically significant in improving max vertical jumps.\n")

} else {

cat("There is no significant improvement in max vertical jumps after the training program.\n")

}
```

Output:

```
Paired t-test

data: before_training and after_training

t = -3.2262, df = 19, p-value = 0.004445

alternative hypothesis: true mean difference is not equal to 0

95 percent confidence interval:
```

-1.5663255 -0.3336745

sample estimates:

mean difference

-0.95

The training program is statistically significant in improving max vertical jumps.

Manual:

Manual:

Null Hypothesis(H_0) : There is no difference in means ($\mu_{before} = \mu_{after}$)

Alternative Hypothesis(H_1) : There is a significant difference in means($\mu_{before} \neq \mu_{after}$)

Player	Max Vertical Jump Before Training Program	Max Vertical Jump After Training Program	Difference (d)	$d - \overline{d}$	$(d - d)^2$
1	22	24	-2	-1.05	1.1025
2	20	22	-2	-1.05	1.1025
3	19	19	0	0.95	0.9025
4	24	22	2	2.95	8.7025
5	25	28	-3	-2.05	4.2025
6	25	26	-1	-0.05	0.0025
7	28	28	0	0.95	0.9025
8	22	24	-2	-1.05	1.1025
9	30	30	0	0.95	0.9025
10	27	29	-2	-1.05	1.1025
11	24	25	-1	-0.05	0.0025
12	18	20	-2	-1.05	1.1025
13	16	17	-1	-0.05	0.0025
14	19	18	1	1.95	3.8025
15	19	18	1	1.95	3.8025
16	28	28	0	0.95	0.9025
17	24	26	-2	-1.05	1.1025
18	25	27	-2	-1.05	1.1025
19	25	27	-2	-1.05	1.1025
20	23	24	-1	-0.05	0.0025
			$\Sigma d = -19$		$\Sigma(d - \overline{d})^2 = 32.95$

$$Sample\ Mean\ of\ Difference,\ \overline{d} = \frac{\Sigma d}{n} = \frac{-19}{20} = -\ 0.95$$

$$Standard\ Deviation\ of\ Difference,\ S_d = \sqrt{\frac{\Sigma(d-\overline{d})^2}{n-1}}$$

$$S_d = \sqrt{\frac{32.95}{19}} = \sqrt{1.73421} = 1.316894$$

$$Standard\ Error\ of\ Mean\ Difference,\ SE(\overline{d}) = \frac{S_d}{\sqrt{n}}$$

$$SE(\overline{d}) = \frac{1.316894}{\sqrt{20}} = \frac{1.316894}{4.472136} = 0.29447$$

$$t - statistic,\ t = \frac{\overline{d}}{SE(d)}$$

$$t = \frac{\overline{d}}{SE(d)} = \frac{-0.95}{0.29447} = -3.2261$$

$$Determine\ Degrees\ of\ Freedom,\ df = n - 1 = 20 - 1 = 19$$

Finding Critical t-value from t table using significance level(0.01) and degree of freedom(19)
t(0.01, 19) = 2.861

t Table

Significance level

Degree of Freedom

cum. prob	$t_{.50}$	$t_{.75}$	$t_{.80}$	$t_{.85}$	$t_{.90}$	$t_{.95}$	$t_{.975}$	$t_{.99}$	$t_{.995}$	$t_{.999}$	$t_{.9995}$
one-tail	0.50	0.25	0.20	0.15	0.10	0.05	0.025	0.01	0.005	0.001	0.0005
two-tails	1.00	0.50	0.40	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0.01	0.002	0.001
df											
1	0.000	1.000	1.376	1.963	3.078	6.314	12.71	31.82	[illegible]	318.31	636.62
2	0.000	0.816	1.061	1.386	1.886	2.920	4.303	6.965	[illegible]	22.327	31.599
3	0.000	0.765	0.978	1.250	1.638	2.353	3.182	4.541	[illegible]	10.215	12.924
4	0.000	0.741	0.941	1.190	1.533	2.132	2.776	3.747	[illegible]	7.173	8.610
5	0.000	0.727	0.920	1.156	1.476	2.015	2.571	3.365	[illegible]	5.893	6.869
6	0.000	0.718	0.906	1.134	1.440	1.943	2.447	3.143	[illegible]	5.208	5.959
7	0.000	0.711	0.896	1.119	1.415	1.895	2.365	2.998	[illegible]	4.785	5.408
8	0.000	0.706	0.889	1.108	1.397	1.860	2.306	2.896	[illegible]	4.501	5.041
9	0.000	0.703	0.883	1.100	1.383	1.833	2.262	2.821	[illegible]	4.297	4.781
[illegible]	0.000	0.700	0.879	1.093	1.372	1.812	2.228	2.764	[illegible]	4.144	4.587
11	0.000	0.697	0.876	1.088	1.363	1.796	2.201	2.718	[illegible]	4.025	4.437
[illegible]	0.000	0.695	0.873	1.083	1.356	1.782	2.179	2.681	[illegible]	3.930	4.318
[illegible]	0.000	0.694	0.870	1.079	1.350	1.771	2.160	2.650	[illegible]	3.852	4.221
[illegible]	0.000	0.692	0.868	1.076	1.345	1.761	2.145	2.624	[illegible]	3.787	4.140
[illegible]	0.000	0.691	0.866	1.074	1.341	1.753	2.131	2.602	[illegible]	3.733	4.073
[illegible]	0.000	0.690	0.865	1.071	1.337	1.746	2.120	2.583	[illegible]	3.686	4.015
[illegible]	0.000	0.689	0.863	1.069	1.333	1.740	2.110	2.567	[illegible]	3.646	3.965
[illegible]	0.000	0.688	0.862	1.067	1.330	1.734	2.101	[illegible]	[illegible]	3.610	3.922
19	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	2.861	3.579	3.883
[illegible]	0.000	0.687	0.860	1.064	1.325	1.725	2.086	[illegible]	[illegible]	3.552	3.850
21	0.000	0.686	0.859	1.063	1.323	1.721	2.080	2.518	2.831	3.527	3.819

absolute value of the calculated t-value is greater than the critical t-value,
So, reject the null hypothesis

The training program is statistically significant in improving max vertical jumps

8. A company has three manufacturing plants, and company officials want to determine whether there is a difference in the average age of workers at the three locations. The following data are the ages of five randomly selected workers at each plant. Perform a one-way ANOVA to determine whether there is a significant difference in the mean ages of the workers at three plants. Use α=0.01. Write an R script for the above problem.

Plant(Employee Ages)

1	2	3
29	32	25
27	33	24
30	31	24
27	34	25
28	30	25

```
plant1 <- c(29, 27, 30, 27, 28)

plant2 <- c(32, 33, 31, 34, 30)

plant3 <- c(25, 24, 24, 25, 25)

data1 <- data.frame(

Plant = factor(rep(1:3, each = 5)),

Age = c(plant1, plant2, plant3)

)

print("Data:")

print(data1)

result <- aov(Age ~ Plant, data = data1)
```

```
print("ANOVA Results:")
summary_result <- summary(result)
print(summary_result)
pvalue <- summary_result[[1]][["Pr(>F)"]][1]
if (pvalue < 0.01) {
cat("There is a significant difference in the mean ages of workers at three plants (p-value =",
pvalue, ")")
} else {
cat("There is no significant difference in the mean ages of workers at three plants (p-value =",
pvalue, ")")
}
```

Output:

```
One-Way ANOVA Result:
Df Sum Sq Mean Sq F value Pr(>F)
Plant 2 136.9 68.47 45.64 2.46e-06 ***
Residuals 12 18.0 1.50
---
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
There is a significant difference in the mean ages of workers at the three plants (p-value =
2.459041e-06 )
```

Manual:

H_0: There is no significant difference in the mean ages of the workers at three plants

$\mu_1 = \mu_2 = \mu_3$

H_1: There is significant difference in the mean ages of the workers at three plants $\mu_1 \neq \mu_2 \neq \mu_3$

$$Mean\ of\ plant\ 1,\ x_1 = \frac{29 + 27 + 30 + 27 + 28}{5} = \frac{141}{5} = 28.2$$

$$Mean\ of\ plant\ 2,\ x_2 = \frac{32 + 33 + 31 + 34 + 30}{5} = \frac{160}{5} = 32$$

$$Mean\ of\ plant\ 3,\ x_3 = \frac{25 + 24 + 24 + 25 + 25}{5} = \frac{123}{5} = 24.6$$

$$Total\ Mean,\ x_t = \frac{x_1 + x_2 + x_3}{3} = \frac{28.2 + 32 + 24.6}{3} = \frac{84.8}{3} = 28.27$$

$$SSBetween = 5.((x_1 - x_t)^2 + (x_2 - x_t)^2 + (x_3 - x_t)^2)$$

$$SSBetween = 5.((28.2 - 28.27)^2 + (32 - 28.27)^2 + (24.6 - 28.27)^2)$$

$$SSBetween = 5.((-0.07)^2 + (3.73)^2 + (-3.67)^2)$$

$$SSBetween = 5.(0.0049 + 13.913 + 13.469)$$

$$SSBetween = 5.27.3869$$

$$SSBetween = 136.9345$$

plant 1	$(plant\ 1 - x_1)^2$	plant 2	$(plant\ 2 - x_2)^2$	plant 3	$(plant\ 3 - x_3)^2$
29	$(29 - 28.2)^2 = 0.64$	32	$(32 - 32)^2 = 0$	25	$(25 - 24.6)^2 = 0.16$
27	$(27 - 28.2)^2 = 1.44$	33	$(33 - 32)^2 = 1$	24	$(24 - 24.6)^2 = 0.36$
30	$(30 - 28.2)^2 = 3.24$	31	$(31 - 32)^2 = 1$	24	$(24 - 24.6)^2 = 0.36$
27	$(27 - 28.2)^2 = 1.44$	34	$(34 - 32)^2 = 4$	25	$(25 - 24.6)^2 = 0.16$
28	$(28 - 28.2)^2 = 0.04$	30	$(30 - 32)^2 = 4$	25	$(25 - 24.6)^2 = 0.16$
	$\Sigma(plant\ 1 - x_1)^2 = 6.8$		$\Sigma(plant\ 2 - x_2)^2 = 10$		$\Sigma(plant\ 3 - x_3)^2 = 1.2$

$$SSWithin = \frac{\Sigma(plant\ 1 - x_1)^2 + \Sigma(plant\ 2 - x_2)^2 + \Sigma(plant\ 3 - x_3)^2}{3}$$

$$SSWithin = \frac{6.8 + 10 + 1.2}{3} = 18$$

$$dfBetween = Number\ of\ groups(plants) - 1 = 3 - 1 = 2$$

$$dfWithin = Total\ number\ of\ observation - number\ of\ groups(plants) = 15 - 3 = 12$$

$$MSBetween = \frac{SSBetween}{dfBetween} = \frac{136.9345}{2} = 68.467$$

$$MSWithin = \frac{SSWithin}{dfWithin} = \frac{18}{12} = 1.5$$

$$F - statistic = \frac{MSBetween}{MSWithin} = \frac{68.467}{1.5} = 45.645$$

Finding Critical F - value from F-Distribution: α = 0.01 table

dfBetween is 2, dfWithin is 12
Use this to find Critical F - value
Use dfBetween as Column and dfWithin as row

Critical Values of the F-Distribution: $\alpha = 0.01$

Denom.	Numerator Degrees of Freedom (dfBetween)									
d.f. (dfWithin)	1	2	3	4	5	6	7	8	9	10
1	4052.181	4999.500	5403.352	5624.583	5763.650	5858.986	5928.356	5981.070	6022.473	6055.847
2	98.503	[illegible]	99.166	99.249	99.299	99.333	99.356	99.374	99.388	99.399
3	34.116	[illegible]	29.457	28.710	28.237	27.911	27.672	27.489	27.345	27.229
4	21.198	[illegible]	16.694	15.977	15.522	15.207	14.976	14.799	14.659	14.546
5	16.258	[illegible]	12.060	11.392	10.967	10.672	10.456	10.289	10.158	10.051
6	13.745	[illegible]	9.780	9.148	8.746	8.466	8.260	8.102	7.976	7.874
7	12.246	[illegible]	8.451	7.847	7.460	7.191	6.993	6.840	6.719	6.620
8	11.259	[illegible]	7.591	7.006	6.632	6.371	6.178	6.029	5.911	5.814
9	10.561	[illegible]	6.992	6.422	6.057	5.802	5.613	5.467	5.351	5.257
10	10.044	[illegible]	6.552	5.994	5.636	5.386	5.200	5.057	4.942	4.849
11	9.646	[illegible]	6.217	5.668	5.316	5.069	4.886	4.744	4.632	4.539
12	[illegible]	$\boxed{6.927}$	5.953	5.412	5.064	4.821	4.640	4.499	4.388	4.296
13	9.074	6.701	5.739	5.205	4.862	4.620	4.441	4.302	4.191	4.100
14	8.862	6.515	5.564	5.035	4.695	4.456	4.278	4.140	4.030	3.939
15	8.683	6.359	5.417	4.893	4.556	4.318	4.142	4.004	3.895	3.805

Critical F-value = 6.927

There is significant difference in the mean ages of the workers at three plants

Tip: If your answer don't match with the output. (Only use as last resort)

You can write a code same as your manual solution like by converting formula and value in manual to formula and variable in code.

For example:

B6th program

Actual code:

```
lower_bound <- 200

upper_bound <- 240

x1 <- 230

probability_x1 <- 1 - punif(x1, lower_bound, upper_bound)

x2_min <- 205

x2_max <- 220

probability_x2 <- punif(x2_max, lower_bound, upper_bound) - punif(x2_min, lower_bound,
upper_bound)

cat("a. P(x>230):", probability_x1, "\n")

cat("b. P(205≤x≤220):", probability_x2, "\n")
```

Code made as same as manual:

```
lower_bound <- 200

upper_bound <- 240

x1 <- 230

probability_x1_manual <- (upper_bound - x1) / (upper_bound - lower_bound)

x2_min <- 205

x2_max <- 220

probability_x2_manual <- (x2_max - x2_min) / (upper_bound - lower_bound)

cat("a. Manual calculation for P(x > 230):", probability_x1_manual, "\n")

cat("b. Manual calculation for P(205 ≤ x ≤ 220):", probability_x2_manual, "\n")
```