

---

Q.1(A)

$$\text{lik}(\theta) = f_1(3) \times f_1(4) \times f_1(1) = \theta^3(1 - \theta)^3 \times \theta^4(1 - \theta)^4 \times \theta^1(1 - \theta)^1 = \theta^8(1 - \theta)^8$$

A = 8 and B = 8

likelihood function is  $\text{lik}(\theta) = \theta^8(1 - \theta)^8$

---

Q.1(B)

MLE estimate using numerical optimization  $\hat{\theta} = 0.5$

---

Q.1(R Commands)

```
> #1(a)
> #Likelihood function
> lik <- function(theta) {
+   theta^8 * (1 - theta)^8
+ }
> #1(b)
> #Maximum Likelihood Estimation (MLE):
> estimate <- optimize(lik, interval = c(0, 1), maximum = TRUE)$maximum
> estimate
[1] 0.5
```

## Q.2

- (A) Find a 95% confidence interval for the variance  $\sigma^2 = (5.538468, 39.01548)$
- (B) Find a 95% confidence interval for the standard deviation  $\sigma = (2.353395, 6.246237)$
- (C) Find a 90% one-sided CI that gives an upper bound on standard deviation  $\sigma = (0, 7.42165)$

## Q.2 R Commands

```
> rm(list=ls())
> #2
> sample <- c(8.58, 5.21, 6.13, 4.41, 0.16, 1.03, 0.36, -0.60, -1.95, 0.87)
> n <- length(sample) # sample size
> xbar <- mean(sample) # sample mean
> s_squared <- var(sample) #sample variance
> # chi-square distribution (two-tailed)
> alpha <- 0.05
> chi_sq <- qchisq(alpha / 2, df = n - 1)
> chi_sq_minus <- qchisq(1 - alpha / 2, df = n - 1)
> # one-sided t-distribution (upper tail)
> alpha_90 <- 0.10
> t_1_minus_alpha <- qt(1 - alpha_90, df = n - 1)
> # 95% confidence interval for variance
> ci_var <- c((n - 1) * s_squared / chi_sq_minus, (n - 1) * s_squared / chi_sq)
> # 95% confidence interval for standard deviation
> ci_sd <- sqrt(ci_var)
> # One-sided 90% CI (upper bound) for standard deviation
> ci_sd_upper <- sqrt((n - 1) * s_squared) / t_1_minus_alpha
> # Print results
> cat("95% CI for variance:", ci_var, "\n")
95% CI for variance: 5.538468 39.01548
> cat("95% CI for standard deviation:", ci_sd, "\n")
95% CI for standard deviation: 2.353395 6.246237
> cat("90% one-sided CI (upper bound) for standard deviation:", c(0, ci_sd_upper), "\n")
90% one-sided CI (upper bound) for standard deviation: 0 7.42165
```

## Q.3

(A) Point estimate  $\hat{p}$  for the success probability  $p$ : 0.2333333

(B) 90% confidence interval for the success probability  $p$ : 0.1063174, 0.3603493

## Q.3 (R commands)

```
> #3
> n <- 30
> Y <- 7
> alpha <- 0.10 # 90% confidence level
> p <- Y / n
> p # Point estimate for p
[1] 0.2333333
> # Standard error
> SE <- sqrt((p * (1 - p)) / n)
> # Critical value (z_alpha/2)
> z <- qnorm(1 - alpha / 2)
> # Confidence interval
> lower_bound <- p - z * SE
> upper_bound <- p + z * SE
> c(lower_bound, upper_bound)
[1] 0.1063174 0.3603493
```

## Q.4

(A)

Bootstrap Confidence Interval (95%) for AML patients: -0.90867, -0.06290364

Bootstrap Confidence Interval (95%) for ALL patients: 0.9777203, 1.888359

(B)

T-Interval (95%) for AML patients: -0.9722676, 0.01140578

T-Interval (95%) for ALL patients: 0.9633854, 1.940744

(C) There appears to be a difference in MB-1 gene expression between AML and ALL patients.

## Q.4 (R Commands )

```
> library(boot)
> library(multtest)
> data(golub)
> gol.fac <- factor(golub.cl, levels = 0:1, labels = c("ALL", "AML"))
> gene_index <- grep("MB-1 gene", golub.gnames[, 2])
> MB1AML <- golub[gene_index, gol.fac == "AML"]
> MB1ALL <- golub[gene_index, gol.fac == "ALL"]
> nAML <- length(MB1AML)
> nALL <- length(MB1ALL)
> nboot <- 1000
> boot_means_AML <- rep(NA, nboot)
> boot_means_ALL <- rep(NA, nboot)
> for (i in 1:nboot) {
+   data_star_AML <- sample(MB1AML, replace = TRUE)
+   boot_means_AML[i] <- mean(data_star_AML)
+
+   data_star_ALL <- sample(MB1ALL, replace = TRUE)
+   boot_means_ALL[i] <- mean(data_star_ALL)
+ }
> boot_ci_AML <- quantile(boot_means_AML, c(0.025, 0.975))
> boot_ci_ALL <- quantile(boot_means_ALL, c(0.025, 0.975))
```

```
> cat("Bootstrap Confidence Interval (95%) for AML patients:", boot_ci_AML, "\n")
Bootstrap Confidence Interval (95%) for AML patients: -0.90867 -0.06290364
> cat("Bootstrap Confidence Interval (95%) for ALL patients:", boot_ci_ALL, "\n")
Bootstrap Confidence Interval (95%) for ALL patients: 0.9777203 1.888359
> #4(b)
> t_CI_AML <- t.test(MB1AML)$conf.int
> t_CI_ALL <- t.test(MB1ALL)$conf.int
> cat("T-Interval (95%) for AML patients:", t_CI_AML, "\n")
T-Interval (95%) for AML patients: -0.9722676 0.01140578
> cat("T-Interval (95%) for ALL patients:", t_CI_ALL, "\n")
T-Interval (95%) for ALL patients: 0.9633854 1.940744
> #4(c)
> cat("\nBased on the confidence intervals:")
```

Based on the confidence intervals:

```
> if (t_CI_AML[1] > t_CI_ALL[2] || t_CI_AML[2] < t_CI_ALL[1]) {
+   cat("There appears to be a difference in MB-1 gene expression between AML and ALL patients.\n")
+ } else {
+   cat("There doesn't appear to be a significant difference in MB-1 gene expression between AML and
ALL patients.\n")
+ }
There appears to be a difference in MB-1 gene expression between AML and ALL patients.
```