

114-1-程式設計與統計軟體 小考(2) 答案卷

系級：_____ 學號：_____ 姓名：_____

注意事項：

1. 本答案卷上請務必於 Console 內複製「執行後的程式碼、結果及圖形」，於本答案卷貼上 (Courier New, 10 點字，白底黑字)，不是只有程式碼，不是只有結果或報表。貼圖時，請注意圖形之大小，應適中；圖內之文字數字應可識別。
2. 作答完請將此 word 檔存檔，檔名為「姓名-StatSoft-exam1.docx」(更改成自己「姓名」)並上傳。
3. 帳號：statsoft，密碼：1141。
4. 如果上傳網站出現空白頁，請將滑鼠移至「網址列」後，按「Enter」即可。若再不行，請清除快取 Cache 再試一次。若再不行，請換其它瀏覽器 (IE/Edge/Firefox/Chrome)
5. 答案卷常見錯誤填答格式提醒 (請務必避免) (已事先於課程 FB 公告)：
https://hmwu.idv.tw/web/teaching/114-1/StatSoftware/R-WrongFormat_remark.pdf

```
# ex1
> department <- "統計二"
> id <- "1234567"
> name <- "吳小銘"
> oath_content <- readLines(file("Data/Oath.txt", encoding = "Big5"), warn = F)
> cat(oath_content[1],": ", department, ",", id, ",", name, "\n", oath_content[2],
sep = "")
```

本人：統計二，1234567，吳小銘

恪遵各項考試規則，若如違反，願受校方最嚴厲處罰，謹誓。

```
# ex2
(a)
> my_dnorm <- function(x, mu, sigma) {
+
+   f <- (1 / (sqrt(2 * pi) * sigma)) * exp(-0.5 * ((x - mu) / sigma)^2)
+   f
+ }
```

```
(b)
> library(ggplot2)
> mu0 <- 0
> mua <- 3.5
> sigma <- 1.5
```

```

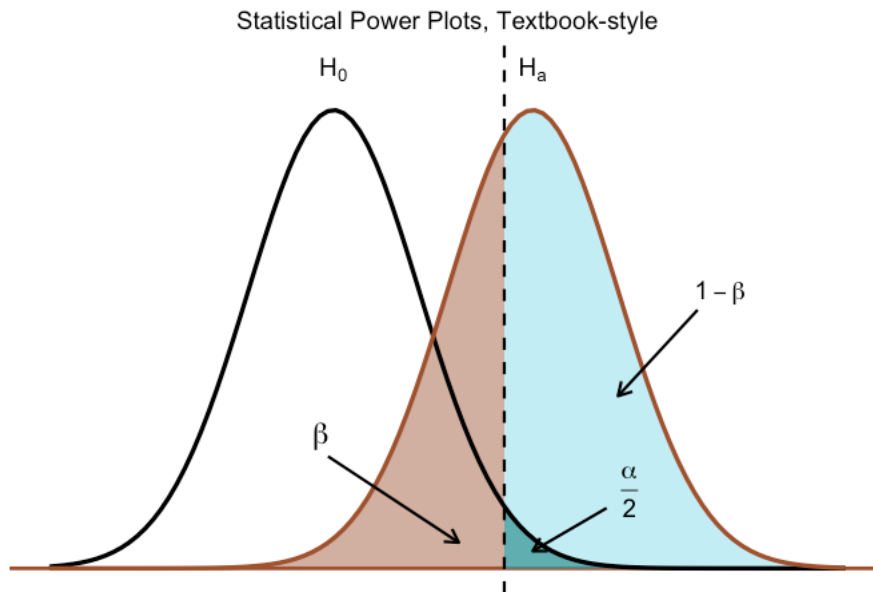
> threshold <- 3
> color_H0_line <- "black"
> color_Ha_line <- "#A0522D"
> fill_beta <- alpha("#A0522D", 0.5)
> fill_power <- alpha("#80DAEB", 0.5)
> fill_alpha <- alpha("#008080", 0.6)
> ggplot(data = data.frame(x = c(-5, 9)), aes(x = x)) +
+
+   theme_classic() +
+   labs(title = "Statistical Power Plots, Textbook-style", x = NULL, y = NULL) +
+   theme(plot.title = element_text(hjust = 0.5, size = 14),
+         axis.line = element_blank(), axis.text = element_blank(),
+         axis.ticks = element_blank()) +
+
+   geom_hline(yintercept = 0, color = color_Ha_line, linewidth = 1) +
+
+   stat_function(fun = my_dnorm, args = list(mu = mua, sigma = sigma),
+               xlim = c(-5, threshold), geom = "area", fill = fill_beta) +
+
+   stat_function(fun = my_dnorm, args = list(mu = mua, sigma = sigma),
+               xlim = c(threshold, 9), geom = "area", fill = fill_power) +
+
+   stat_function(fun = my_dnorm, args = list(mu = mu0, sigma = sigma),
+               xlim = c(threshold, 9), geom = "area", fill = fill_alpha) +
+
+   geom_vline(xintercept = threshold, linetype = "dashed", linewidth = 0.8) +
+
+   stat_function(fun = my_dnorm, args = list(mu = mu0, sigma = sigma),
+               linewidth = 1.2, color = color_H0_line) +
+
+   stat_function(fun = my_dnorm, args = list(mu = mua, sigma = sigma),
+               linewidth = 1.2, color = color_Ha_line) +
+
+   annotate("text", x = mu0, y = 0.29, label = "H[0]", parse = TRUE, size = 5) +
+   annotate("text", x = mua, y = 0.29, label = "H[a]", parse = TRUE, size = 5) +
+
+   annotate("text", x = 6.8, y = 0.16, label = "1 - beta", parse = TRUE, size = 5) +
+   annotate("segment", x = 6.4, y = 0.15, xend = 5.0, yend = 0.1,
+         arrow = arrow(length = unit(0.3, "cm")), linewidth = 0.8) +
+

```

```

+ annotate("text", x = -0.22, y = 0.077, label = "beta", parse = TRUE, size = 6) +
+ annotate("segment", x = -0.1, y = 0.065, xend = 2.2, yend = 0.015,
+       arrow = arrow(length = unit(0.3, "cm")), linewidth = 0.8) +
+
+ annotate("text", x = 5.2, y = 0.045, label = "frac(alpha, 2)", parse = TRUE, size =
5) +
+ annotate("segment", x = 4.8, y = 0.035, xend = 3.5, yend = 0.01,
+       arrow = arrow(length = unit(0.3, "cm")), linewidth = 0.8)

```



```

# ex3
(a)
> sales_ds <- read.csv("Data/AppleStore_Sales.txt", header = TRUE, sep="\t")
> cols_quantity <- paste("Quantity", stores, sep = "_")
> cols_price <- paste("UnitPrice", stores, sep = "_")
> cols_sales <- paste("Sales", stores, sep = "_")
> stacked_quantity <- stack(sales_ds[, cols_quantity])
> stacked_price <- stack(sales_ds[, cols_price])
> stacked_sales <- stack(sales_ds[, cols_sales])
> num_stores <- length(stores)
> id_data <- sales_ds[rep(1:nrow(sales_ds), times = num_stores), c("Date", "Item")]
> store_names <- sub("Quantity_", "", stacked_quantity$ind)
> tidy_data <- data.frame(
+   Date = id_data$Date,
+   Store = store_names,
+   Item = id_data$Item,
+   Quantity = stacked_quantity$values,

```

```

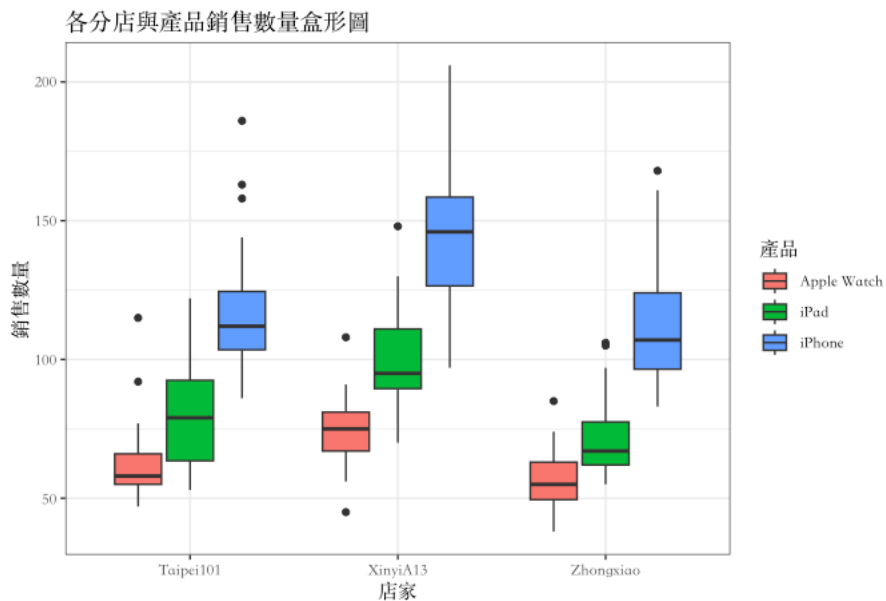
+ UnitPrice = stacked_price$values,
+ Sales = stacked_sales$values
+ )
> rownames(tidy_data) <- NULL
> head(tidy_data)
  Date      Store      Item Quantity UnitPrice  Sales
1 2025/1/1 XinyiA13   iPhone     129    32000 4128000
2 2025/1/1 XinyiA13   iPad       90    16000 1440000
3 2025/1/1 XinyiA13 Apple Watch    67    12000  804000
4 2025/1/2 XinyiA13   iPhone     138    32000 4416000
5 2025/1/2 XinyiA13   iPad      115    16000 1840000
6 2025/1/2 XinyiA13 Apple Watch    59    12000  708000

```

```

(b)
> ggplot(tidy_data, aes(x = Store, y = Quantity, fill = Item)) +
+   geom_boxplot() +
+   labs(title = "各分店與產品之銷售數量分佈", x = "店家", y = "銷售數量", fill = "產品") +
+   theme_bw() +
+   theme(text = element_text(family = "Songti TC"))

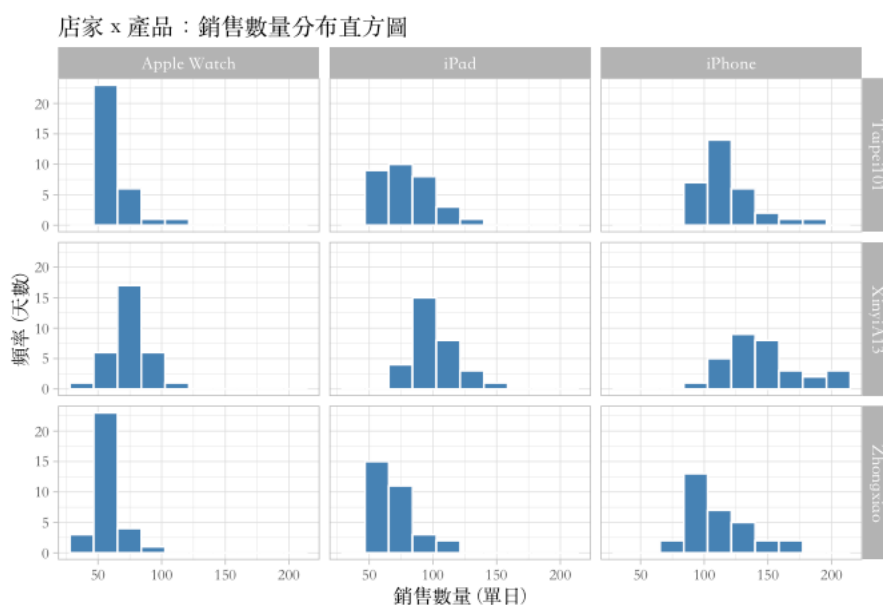
```



```

(c)
> ggplot(tidy_data, aes(x = Quantity)) +
+
+   geom_histogram(fill = "steelblue", bins = 10, color = "white") +
+
+   facet_grid(Store ~ Item) +
+
+   labs(title = "店家 x 產品：銷售數量分布直方圖", x = "銷售數量 (單日)", y = "頻率 (天數)") +
+
+   theme_light() +
+
+   theme(text = element_text(family = "Songti TC"), strip.text = element_text(size =
10))

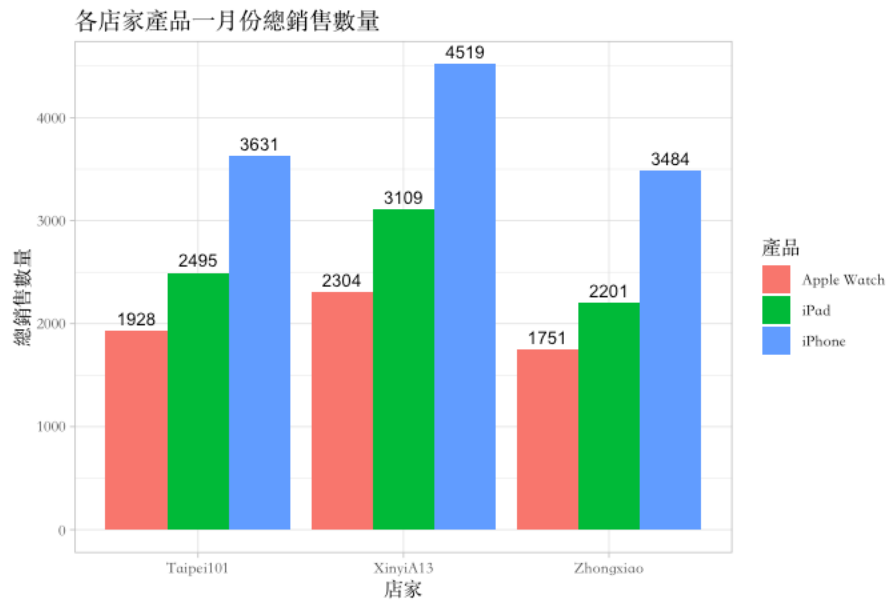
```



```

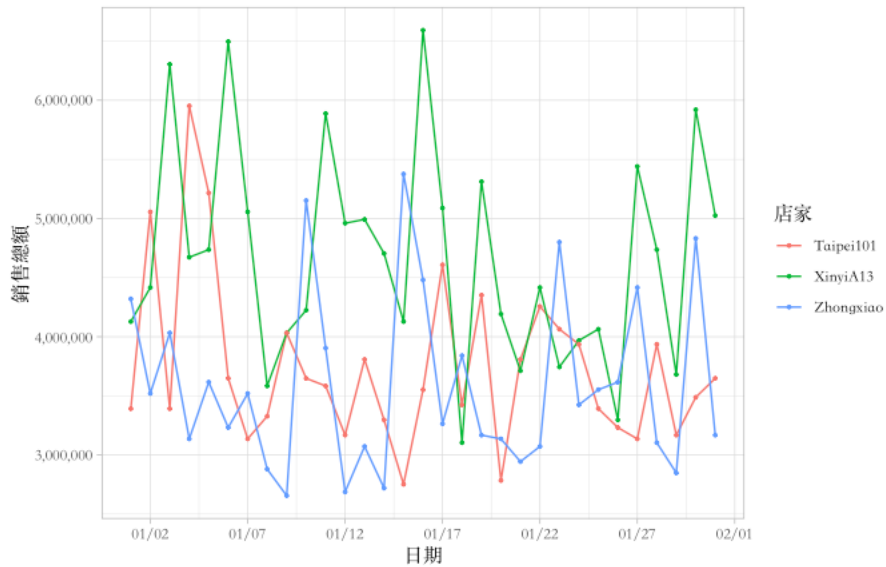
(d)
> plot_ds <- aggregate(Quantity ~ Store + Item, data = tidy_data, FUN = sum)
> ggplot(plot_ds, aes(x = Store, y = Quantity, fill = Item)) +
+   geom_col(position = "dodge") +
+
+   labs(title = "各店家產品一月份總銷售數量", x = "店家", y = "總銷售數量", fill = "產品") +
+
+   geom_text(aes(label = Quantity), position = position_dodge(width = 0.9), vjust = -
0.5, size = 3.5) +
+
+   theme_light() +
+
+   theme(text = element_text(family = "Songti TC"))

```



```
(e)
> iphone_data <- subset(tidy_data, Item == "iPhone")
> iphone_data$Date <- as.Date(iphone_data$Date, format = "%Y/%m/%d")
> ggplot(iphone_data, aes(x = Date, y = Sales, color = Store)) +
+   geom_line(linewidth = 0.5) +
+
+   geom_point(size = 0.6) +
+
+   labs(title = "三家店家每日 iPhone 銷售總額時間序列圖", x = "日期", y = "銷售總額", color =
"店家" ) +
+
+   scale_x_date(date_labels = "%m/%d", date_breaks = "5 days") +
+
+   scale_y_continuous(labels = scales::comma) +
+
+   theme_light() +
+
+   theme(text = element_text(family = "Songti TC"))
```

三家店家每日 iPhone 銷售總額時間序列圖



```
(f)
> iphone_df <- subset(tidy_data, Item == "iPhone", select = c("Date", "Store",
"Quantity"))
> ipad_df <- subset(tidy_data, Item == "iPad", select = c("Date", "Store",
"Quantity"))
> scatter_ds <- merge(iphone_df, ipad_df, by = c("Date", "Store"))
> colnames(scatter_ds) <- c("Date", "Store", "iPhone_quantity", "iPad_quantity")
> ggplot(scatter_ds, aes(x = iPhone_quantity, y = iPad_quantity, color = Store, shape =
Store)) +
+
+   geom_point(size = 3, alpha = 0.8) +
+
+   labs(title = "iPhone 與 iPad 銷售數量散佈圖", x = "iPhone 銷售數量", y = "iPad 銷售數量",
color = "店家", shape = "店家" ) +
+
+   theme_light() +
+
+   theme(text = element_text(family = "Songti TC"))
```

iPhone 與 iPad 銷售數量散佈圖

