107-1 Statistics

LAB5: SAMPLING DISTRIBUTIONS AND PARAMETERS

助教：廖皓宇、吳家禎、賴冠宇

2018/10/26

THIS WEEK LAB

1026實習：樣本分布與統計參數

1. 利用電腦模擬，建立以下範例的樣本分佈直方圖，以此樣本分佈，計算其平均數與標準差，並和理論值進行比較。
   1) Example 9.4: Sample Proportions
   2) Example 9.6: Difference in Two Sample Proportions
   3) Example 9.8: Sample Means
   4) Example 9.9: Sample Mean of Paired Differences
   5) Example 9.10: Difference in Two Sample Means
2. 將上題的樣本分佈圖與理論分佈相互疊合與比較。
3. 計算第1題各小題的z-score 。

1) EXAMPLE 9.4: Sample proportions

① **Sampling one time**

```
# Generate random variables (phat: proportion)
bino.x = function(n, p) {

  pp = p*10
  samp = sample(1:10, n, replace = T)

  response = c()
  for (i in 1:length(samp)) {
    if (samp[i] <= pp) {
      response[i] = 1
    } else {
      response[i] = 0
    }
  }

  #another way to write
  samp = sample(x = c(0,1), n, replace = T, prob = c(1-p,p))
  response = samp
  #

  x = sum(response)
  phat = x/n
  return(phat)
}
```

② **Sampling multiple times → sample distribution**

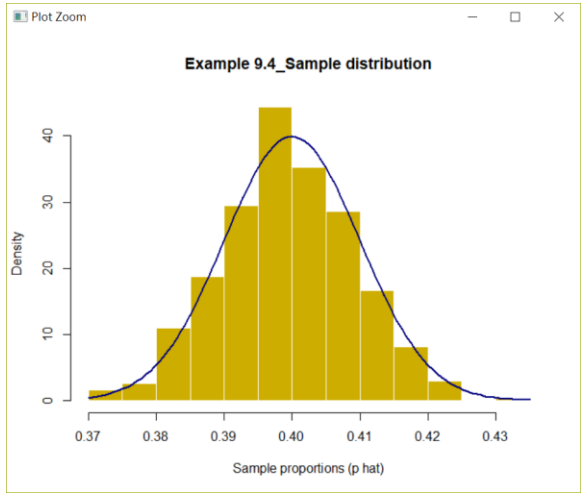
```
# Sampling multiple times (for sample distribution)
simu.phat = c()
for (j in 1:1000) {
  x = bino.x(n = 2400, p = 0.4)
  simu.phat[j] = x
}
```

③ Draw sample distribution

```
# Draw the sampling distribution
hist(simu.phat, probability = T, breaks = 20,
     main = "Example 9.4_Sample distribution", xlab = "Sample proportions (p hat)",
     col = "gold3", border = "white")
```

- **Calculate "mean of proportions" and "sd of proportions"**

```
# calculate the simulated "mean of proportions" and "sd of proportions"
simu.mean = mean(simu.phat); simu.mean
simu.sd = sd(simu.phat); simu.sd
```




```
# calculate theoretical "mean of proportions" and "sd of proportions"
n = 2400
p = 0.4
theo.mean = p; theo.mean
theo.sd = sqrt(p*(1-p)/n); theo.sd
```

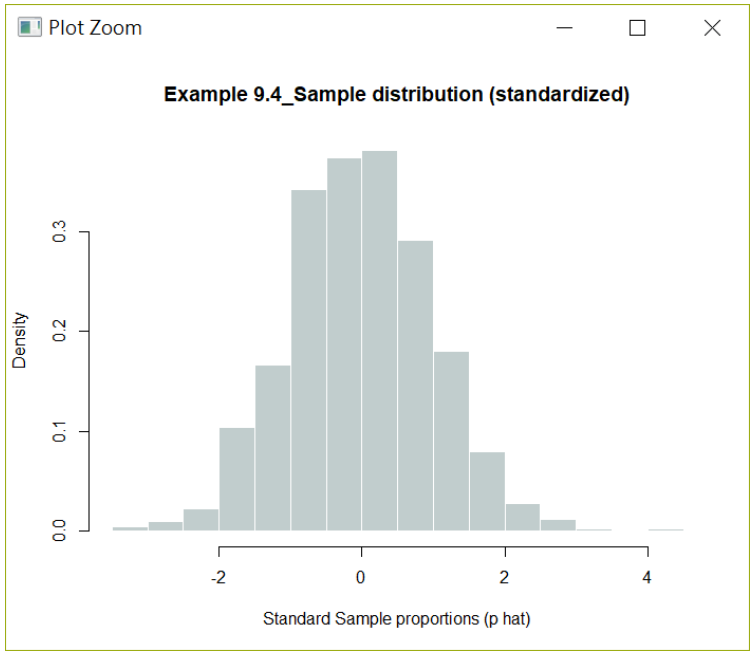
- **Overlapping sample distribution from theoretical results**

```
# Overlapping theoretical results
curve(dnorm(x, mean = theo.mean, sd = theo.sd), add = T, col = "navy", lwd = 2)
```

- **Calculate z score**

```
# Calculating z-score
simu.z = (simu.phat - theo.mean) / (theo.sd)

#distribution of z-score
hist(simu.z, probability = T, breaks = 20,
     main = "Example 9.4_Sample distribution (standardized)", xlab = "Standard Sample proportions (p hat)",
     col = "azure3", border = "white")
```



3) EXAMPLE 9.8: Sample means

① **Sampling one time**

```
# Generate random variables (samp: each observation; xbar = mean of samples)
norm.x = function(mu, sigma, n) {
  samp = rnorm(n, mean = mu, sd = sigma)
  xbar = mean(samp)

  return(xbar)
}
```

② **Sampling multiple times → sample distribution**

```
# Sampling multiple times (for sample distribution)
simu.xbar = c()
for (j in 1:1000) {
  x = norm.x(mu = 8, sigma = 5, n = 25)
  simu.xbar[j] = x
}
```

③ Draw sample distribution

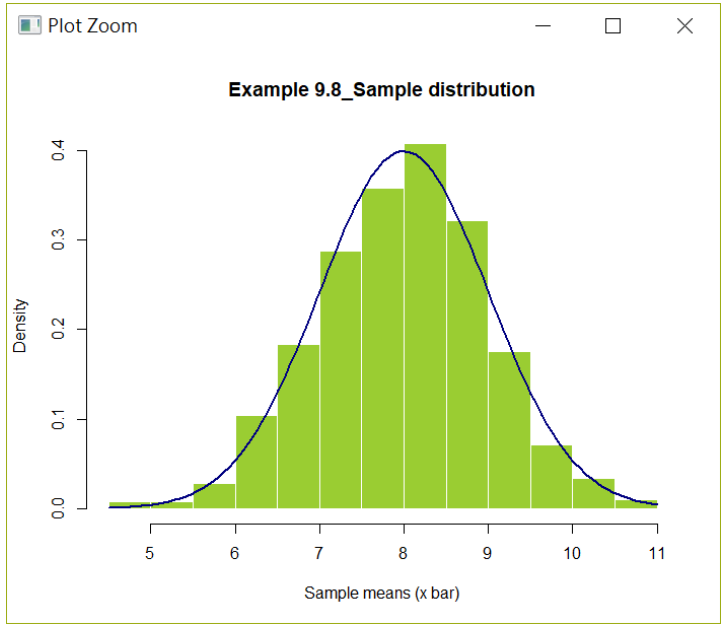
```
# Draw the sampling distribution
hist(simu.xbar, probability = T, breaks = 20,
     main = "Example 9.8_Sample distribution", xlab = "Sample means (x bar)",
     col = "olivedrab3", border = "white")
```

● Calculate "mean of means" and "sd of means"

```
# calculate the simulated "mean of means" and "sd of means"
simu.mean = mean(simu.xbar); simu.mean
simu.sd = sd(simu.xbar); simu.sd


# calculate theoretical "mean of means" and "sd of means"
mu = 8
sigma = 5
n = 25

theo.mean = mu; theo.mean
theo.sd = sigma/sqrt(n); theo.sd
```



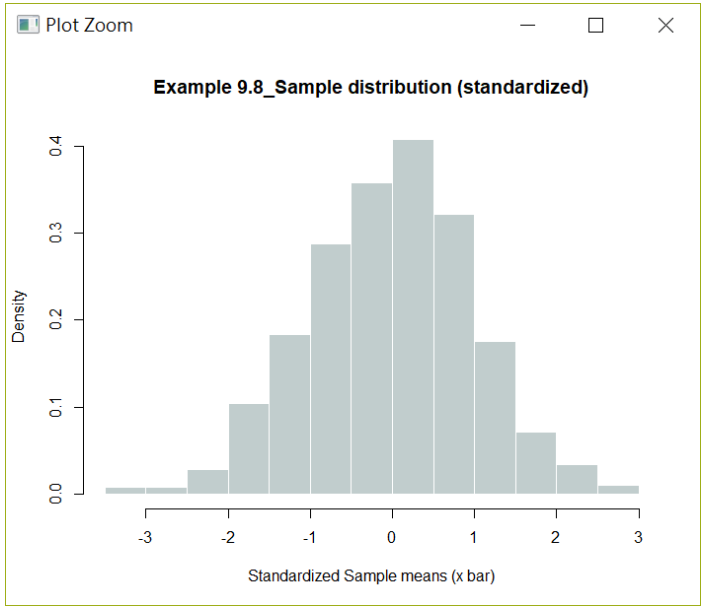
● Overlapping sample distribution from theoretical results

```
# Overlapping theoretical results
curve(dnorm(x, mean = theo.mean, sd = theo.sd), add = T, col = "navy", lwd = 2)
```

- **Calculate z score**

```
# Calculating z-score
simu.z = (simu.xbar - theo.mean) / (theo.sd)

#distribution of z-score
hist(simu.z, probability = T, breaks = 20,
     main = "Example 9.8_Sample distribution (standardized)", xlab = "Standardized Sample means (x bar)",
     col = "azure3", border = "white")
```



本週實習作業

1. 電腦模擬，計算 **模擬的平均數與標準差** 及 **理論的平均值與標準差**，兩者比較。
2. **電腦模擬的樣本分布圖** 與 **理論的樣本分布圖** 疊合、比較。
3. 計算第1題中，**各小題**的z-score。