STAT 542: Homework 1

Arka Mitra (arkam2)

Due: Tuesday 11:59 PM CT, Feb 2nd

Contents

Basic calculus

1. Calculate the derivative of $f(x)$

   (a) $f(x) = e^x$

   Ans: e^x

   (b) $f(x) = \log(1 + x)$

   Ans: $\frac{1}{1+x}$

   (c) $f(x) = \log(1 + e^x)$

   Ans $\frac{e^x}{1+e^x}$

2. Taylor expansion. Let f: $\mathbb{R} \rightarrow \mathbb{R}$ be a twice differentiable function. Please write down the first three terms of its Taylor expansion at point $x = 1$.

   Ans : Taylor Expansion of $f(x)$ at $x = 1$ is $f(x)|_{x=1} = f(1) + \frac{f'(1)}{1!}(x - 1)$

3. For the infinite sum $\sum_{n=1}^{\infty} \frac{1}{n^{\alpha}}$, where α is a positive real number, give the exact range of α such that the series converges.

   Ans : The series converges if $\alpha > 1$

Linear algebra

1. What is the eigendecomposition of a real symmetric matrix $A_{n\times n}$? Write down one form of that decomposition and explain each term in your formula. Based on these terms, derive $A^{-1/2}$.

   Ans : The eigendecomposition of $A_{n\times n}$ is $A = Q\Lambda Q^T$, where Q is a matrix where each column is an eigenvector of A, and, Λ is a matrix whose entries are the eigenvalues of A.

   $A^{-1/2} = (Q\Lambda Q^T)^{-1/2} = (Q^{-1})^{1/2}\Lambda^{-1/2}((Q^T)^{-1})^{1/2}$

   $= (Q^T)^{1/2}\Lambda^{-1/2}(Q^{-1})^{-1/2} = (Q^T)^{1/2}\Lambda^{-1/2}Q^{1/2}$

   [$Q^T = Q^{-1}$, eigenvectors of symmetric matrix are orthogonal]

2. What is a symmetric positive definite matrix $A_{n\times n}$? Give one of equivalent definitions and explain your notation.

   Ans : $A_{n\times n}$ is a positive definite matrix if $Q\Lambda Q^T > 0$, where Q is a matrix where each column is an eigenvector of A, and, Λ is a matrix whose entries are the eigenvalues of A.

3. True/False. If you claim a statement is false, explain why. For two real matrices $A_{m\times n}$ and $B_{n\times m}$

   (a) Rank$(A) = \max\{m, n\}$

   Ans : False. Rank$(A) \leq \min\{m, n\}$

   (b) If $m = n$, then trace$(A) = \sum_{i=1}^{n} A_{ii}$

   Ans : True

   (c) If A is a symmetric matrix, then all eigenvalues of A are real

   Ans: True

   (d) If A is a symmetric matrix, λ_1 and λ_2 are two of its eigen-values (not necessarily different) and v_1,v_2 are the corresponding eigen-vectors, then $v_1^T v_2 = 0$.

   Ans: True

   (e) trace(ABAB) = trace(AABB)

   Ans : False. trace(ABAB) $\leq$ trace(AABB)

Statistics

1. X_1, X_2, ..., X_n are i.i.d. $\mathcal{N}(\mu, \sigma^2)$ random variables, where $\mu \in \mathbb{R}$ and $\sigma > 0$ is finite. Let $\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$.

   (a) What is an unbiased estimator? Is $\bar{X}_n$ an unbiased estimator of μ?

   Ans: An estimator with zero bias between the true and expected value of a statistics is an unbiased estimator.

   Yes, $\bar{X}_n$ an unbiased estimator of μ

   (b) What is $E[(\bar{X}_n)^2]$ in terms of n, μ, σ?

   Ans: We know, $E[\bar{X}_n] = \mu$ and $E[(\bar{X}_n - \mu)^2] = \sigma^2$

   $E[(\bar{X}_n)^2] = E[(\bar{X}_n - \mu + \mu)^2] = E[(\bar{X}_n - \mu)^2] - 2E[(\bar{X}_n - \mu)\mu] + E(\mu^2) = \sigma^2 - 2\mu E(\bar{X}_n - \mu) + \mu^2 = \sigma^2 + \mu^2$

   (c) Give an unbiased estimator of σ^2.

   Ans : An unbiased estimator of σ^2 is $(\bar{X}_n - \mu)^2$ or the variance.

   (d) What is a consistent estimator? Is $\bar{X}_n$ a consistent estimator of μ?

   Ans : An estimator is said to be consistent if given an increase in sample size, the estimator converges to the true value of the statistics of the distribution that is being measured.

   Yes, the sample mean is a consistent estimator.

2. Suppose $X_{p\times 1}$ is a vector of covariates, $\beta_{p\times 1}$ is a vector of unknown parameters, ϵ is the unobserved random noise and we assume the linear model relationship $y = X^T\beta + \epsilon$. Suppose we have n i.i.d. samples from this linear model, and the observed data can be written using the matrix form: $\mathbf{y}_{n\times 1} = \mathbf{X}_{n\times p}\beta_{p\times 1} + \boldsymbol{\epsilon}_{n\times 1}$.

   (a) If we want estimate the unknown β using a least square method, what is the objective function $L(\beta)$ to obtain $\widehat{\beta}$?

   Ans: The objective function is given by $L(\beta) = (\mathbf{y} - \mathbf{X}\beta - \epsilon)^T(\mathbf{y} - \mathbf{X}\beta - \epsilon)$

   (b) What is the solution of $\widehat{\beta}$? Represent the solution using the observed data $\mathbf{y}$ and $\mathbf{X}_{n\times p}$. Note that you may assume that $\mathbf{X}^T\mathbf{X}$ is invertible.

   Ans: Differentiating $L(\beta)$ wrt β, and setting the solution equal to zero, we get $\widehat{\beta} = (X^TX)^{-1}X^T\mathbf{y}$

Programming

1. Use the following code to generate a set of observations $\mathbf{y}$ and $\mathbf{X}_{n\times p}$. Following the previously established formula, Write your own code, instead of using existing functions such as `lm()`, to solve for the least square estimator $\widehat{\beta}$. If you are asked to add an intercept term β_0 into your estimation (even the true $\beta_0 = 0$ in our data generator), what should you do?

```
set.seed(1)
n = 100; p = 5
X = matrix(rnorm(n * p), n, p)
y = X %*% c(1, 0, 0, 1, -1) + rnorm(n)
betahat = solve(t(X) %*% X) %*% t(X) %*% y
betahat
```

```
##              [,1]
## [1,]  0.88369785
## [2,] -0.01725878
## [3,] -0.02658416
## [4,]  0.90233834
## [5,] -1.05639331
```

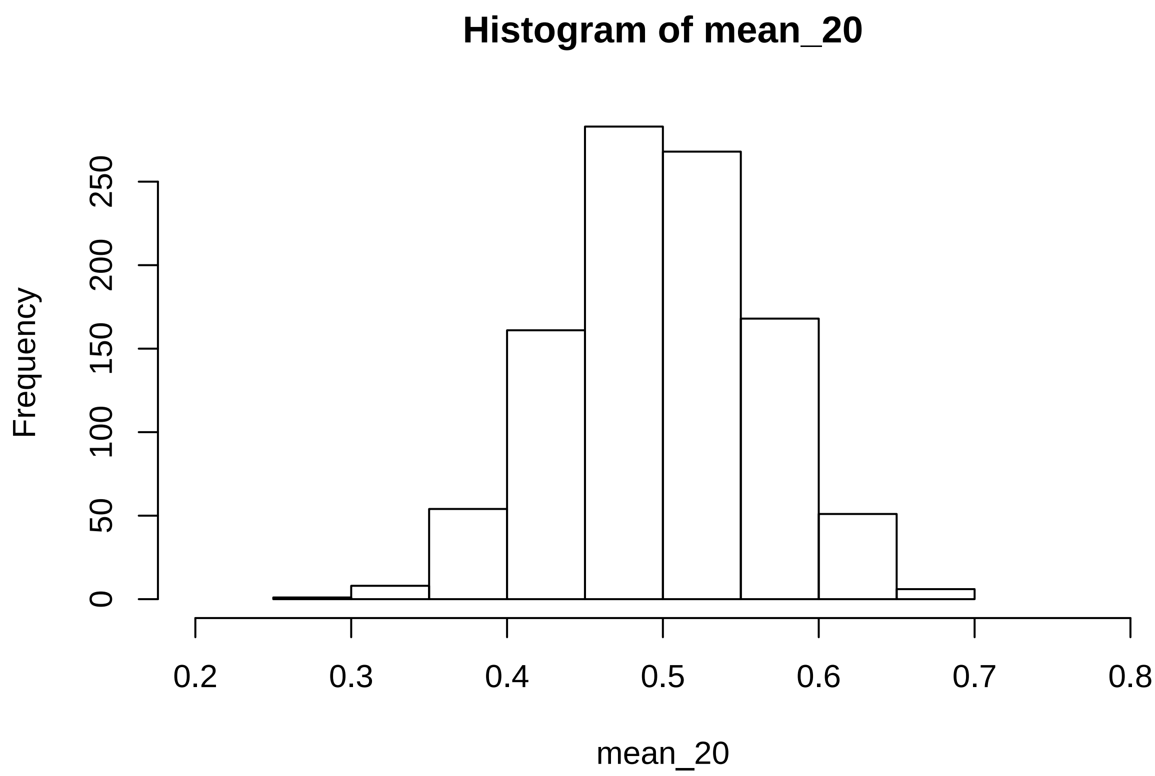
Ans: No change in our estimate of $\widehat{\beta}$.

2. Perform a simulation study to check the consistency of the sample mean estimator $\bar{X}_n$. Please save your random seed so that the results can be replicated by others.

   (a) Generate a set of $n = 20$ i.i.d. observations from uniform (0, 1) distribution and calculate the sample mean $\bar{X}_n$
   (b) Repeat step (a) 1000 times to collect 1000 such sample means and plot them using a histogram.
   (c) How many of such sample means (out of 1000) are at least 0.1 away from true mean parameter, which is 0.5 for uniform (0, 1)?
   (d) Repeat steps (a) to (c) with $n = 100$ and $n = 500$. What conclusion can you make?

```
mean_20 = c(1:1000)*0; mean_100 = c(1:1000)*0; mean_500 = c(1:1000)*0
c_20 = 0; c_100 = 0; c_500 = 0
set.seed(1); oldseed = .Random.seed
for (i in 1:1000) {
```

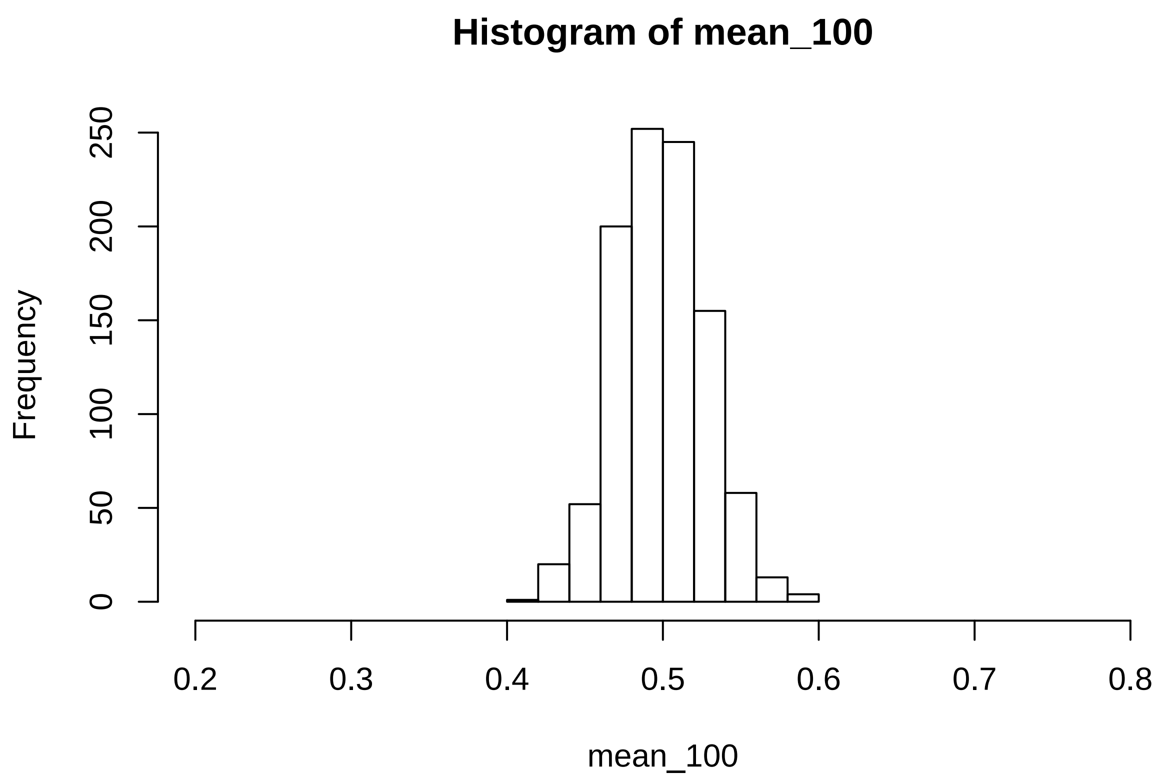
```r
    mean_20[i] = mean(runif(20)); if(abs(mean_20[i]-0.5)>=0.1) c_20=c_20+1
    mean_100[i] = mean(runif(100)); if(abs(mean_100[i]-0.5)>=0.1) c_100=c_100+1
    mean_500[i] = mean(runif(500)); if(abs(mean_500[i]-0.5)>=0.1) c_500=c_500+1
  }
hist(mean_20, breaks=10, xlim=c(0.2,0.8))  # first histogram
```




```r
cat("Number of distant points = ", c_20, "\n")
```

```
## Number of distant points =  120
```

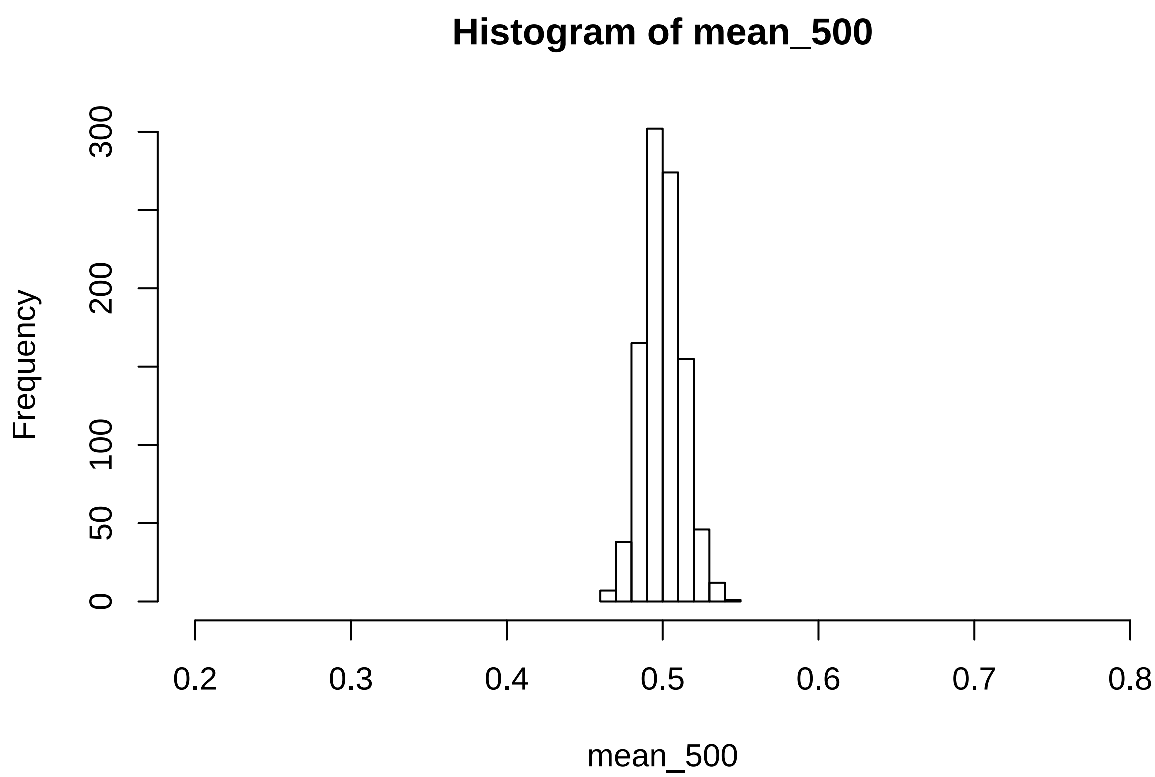
```r
hist(mean_100, breaks=10, xlim=c(0.2,0.8)) # second histogram
```




```
cat("Number of distant points = ", c_100, "\n")
```

```
## Number of distant points =  0
```

```
hist(mean_500, breaks=10, xlim=c(0.2,0.8))  # second histogram
```




```
cat("Number of distant points = ", c_500, "\n")
```

```
## Number of distant points =  0
```

As the number of samples is increased, the mean of the samples closely resembles the true mean of the uniform distribution. Thus, sample mean is a consistent estimator.