1. Please list the graduate level statistics/machine learning/data mining/computer science/econometrics courses you have taken. Include the names of the course with the course number if possible.

2. (Maximum likelihood) Let $X_1, \ldots, X_n$ be iid Gaussian random variables each with probability density function
$$f_X(x; s) = \frac{1}{\sqrt{2\pi s}} \exp\left\{-\frac{x^2}{2s}\right\}.$$
Find the maximum likelihood estimator for s.

3. Let $\widehat{\theta}$ be an estimator for θ. Write the formula for the bias of $\widehat{\theta}$. Is the MLE in the previous problem unbiased? Is it asymptotically unbiased?

4. (Risk analysis) Explain why we might prefer a biased estimator to an unbiased estimator.

5. (Linear methods) Let
$$Y_i = X_i^\top \beta + \epsilon_i, \qquad i = 1, \ldots, n.,$$
where the matrix $\mathbf{X} = [X_1^\top, \ldots, X_n^\top]^\top$ has full rank. Find the ordinary least squares estimator of β.

6. (Coding) What does the following `R` code do?

```
x = 2*pi*runif(100)
y = sin(x) + rnorm(100)
B = splines::bs(x,20)
logSeq <- function(from=1e6, to=1, len=100){
  exp(seq(log(from),log(to),length.out = len))
}
lam = logSeq()
mod = MASS::lm.ridge(y~B, lambda=lam)
preds = cbind(1,B) %*% coef(mod)[which.min(mod$GCV),]
df = data.frame(y=y, x=x, truth=sin(x), preds=preds)
library(tidyverse)
df %>% gather(key="key",value="value",-x) %>%
  ggplot(aes(x=x,y=value,col=key)) + geom_line() + ylab('') +
  theme_minimal()
```

7. (Convex optimization) Write down the Langrage Dual for the following optimization problem and explain the necessary conditions for the stationary point to be an optimum (you do not need to find the stationary point or perform any calculations):
$$\begin{aligned} &\min_b L(y, \mathbb{X}b) \\ \text{subject to} \quad &\|Ab\|^2 \leq C \end{aligned}$$
for some convex function $L : \mathbb{R}^n \to \mathbb{R}^+$, vectors $y \in \mathbb{R}^n$ and $b \in \mathbb{R}^p$, matrices $A \in \mathbb{R}^{m \times p}$ and $\mathbf{X} \in \mathbb{R}^{n \times p}$, and constant C. Here $\|\cdot\|^2$ is the squared ℓ_2-norm of a vector. That is $\|z\|^2 = z^\top z$.