More Problems

This is to prepare you for the final. The difficulty will be slightly above the actual exam. Make sure you can do every problem independently.

1. Consider i.i.d. observations $X_1, ..., X_n \sim N(\mu, 1)$.
   (a) Compute $\mathbb{E}\left(X_1 | \frac{X_1+X_2}{2}\right)$.
   (b) Compute $\mathbb{E}\left(X_1 | \frac{X_1+X_2+X_3+X_4}{4}\right)$.
   (c) Compute $\mathbb{E}\left(\frac{X_1+X_2}{2} \middle| \frac{X_1+X_2+\cdots+X_n}{n}\right)$.
   (d) Compute $\mathbb{E}\left(\frac{X_1+X_2+\cdots+X_n}{n} \middle| \frac{X_1+X_2}{2}\right)$.
   (e) In general, compute $\mathbb{E}\left(\frac{X_1+X_2+\cdots+X_m}{m} \middle| \frac{X_1+X_2+\cdots+X_k}{k}\right)$. You may consider three situations: $m < k$, $m = k$, $m > k$.
   (f) Discuss your finding.
2. Consider independent observations $X_i \sim N(\mu, 2^i)$ for $i = 1, ..., n$.
   (a) Compute the mean squared error $\mathbb{E}(\bar{X} - \mu)^2$, where $\bar{X} = \frac{1}{n}\sum_{i=1}^n X_i$.
   (b) Find the MLE $\hat{\mu}$, and compute $\mathbb{E}(\hat{\mu} - \mu)^2$.
   (c) Which estimator is more accurate?
   (d) Construct a 95% confidence interval using the MLE.
3. Consider i.i.d. observations $X_1, ..., X_n \sim \text{Bernoulli}(p)$.
   (a) Find the MLE of p, denoted by $\hat{p}$.
   (b) Find the asymptotic distribution that $\sqrt{n}(\hat{p} - p)$ converges to.
   (c) Find a variance stabilization transformation so that $\sqrt{n}(g(\hat{p}) - g(p))$ converges to a distribution that does not depend on p.
   (d) Find a testing procedure using $g(\hat{p})$ for $H_0 : p = p_0$ against $H_1 : p > p_0$ with (approximate) significance level 0.05.
4. Consider independent random variables X and Y, with density functions
$$f(x) = \frac{1}{\sqrt{2\pi\sigma_1^2}} e^{-\frac{(x-\mu_1)^2}{2\sigma_1^2}}, \quad g(y) = \frac{1}{\sqrt{2\pi\sigma_2^2}} e^{-\frac{(y-\mu_2)^2}{2\sigma_2^2}}.$$
   Find the density function of $X + Y$. Check whether you can get the density of $N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$.
5. Suppose the density function of X is $f(x)$, what is the density function of $aX + b$?
6. *(Regression with different variances.)* Consider $y_i \sim N(\beta_0 + \beta_1 x_i, \sigma_i^2)$ independently for each $i = 1, ..., n$. Each data point has its own variance σ_i^2, which is assumed known.

(a) Find the MLE of β_0, β_1, denoted as $\hat{\beta}_0, \hat{\beta}_1$.

(b) Find the distribution of $\hat{\beta}_0, \hat{\beta}_1$.

(c) Construction a testing procedure for $H_0 : \beta_1 = 0$ vs $H_1 : \beta \neq 0$.

7. *(Regression Fit.)* Consider the linear model $y \sim N(X\beta, \sigma^2 I_n)$. Let $\hat{\beta}$ be the LSE and $\hat{y} = X\hat{\beta}$ be the fit.

   (a) Find the distribution of $\hat{y}$.

   (b) Suppose you only have $y_1 \sim N(\beta_0 + \beta_1 x_1, \sigma^2)$ and $y_2 \sim N(\beta_0 + \beta_1 x_2, \sigma^2)$. In other words, $p = 2$ and $n = 2$. What is the joint distribution of $\hat{y}_1$ and $\hat{y}_2$?

   (c) Find $\mathbb{E}(\hat{y}_1 | \hat{y}_2)$.

8. *(Common variance, different mean.)* Consider $X_1, ..., X_n \sim N(\mu_1, \sigma^2)$ and $X_{n+1}, ..., X_{2n} \sim N(\mu_2, \sigma^2)$. Everything is independent here.

   (a) Find the MLE of σ^2, denoted as $\hat{\sigma}^2$.

   (b) Based on $\hat{\sigma}^2$, construct an exact confidence interval of σ^2.

   (c) Calculate $\mathbb{E}(\hat{\sigma}^2 - \sigma^2)^2$.

   (d) Find a c such that $\mathbb{E}(c\hat{\sigma}^2 - \sigma^2)^2$ is minimized. Is $c\hat{\sigma}^2$ a better estimator of σ^2 than $\hat{\sigma}^2$.

9. *(Poisson Regression.)* Consider $y_i \sim \text{Poisson}(\beta_1 x_i)$ independently for $i = 1, ..., n$.

   (a) What is the distribution of $\sum_{i=1}^n y_i$?

   (b) Consider the estimator $\hat{\beta}_1 = \bar{y}/\bar{x}$. Find $\mathbb{E}(\hat{\beta}_1)$ and $\mathsf{Var}(\hat{\beta}_1)$.

   (c) Find the asymptotic distribution of $\sqrt{n}(\hat{\beta}_1 - \beta_1)$. (Hint: it is normal, just find the mean and variance.)

   (d) Find a transformation g such that the asymptotic distribution of $\sqrt{n}(g(\hat{\beta}_1) - g(\beta_1))$ does not depend on β_1.

   (e) Construct a testing procedure with significance level (approximately) α for $H_0 : \beta_1 = 0$ vs $H_1 : \beta_1 \neq 0$.

   (f) Is $\hat{\beta}_1$ the MLE?

10. *(Quadratic Regression.)* Consider $y_i = \beta_1 x_i + \beta_2 x_i^2 + e_i$ with $e_i \sim N(0, \sigma^2)$ independently for $i = 1, ..., n$.

    (a) Find the MLE of β_1, β_2, denoted as $\hat{\beta}_1, \hat{\beta}_2$.

    (b) Find the mean and variance of the MLE.

    (c) Construct a t-statistic for testing $H_0 : \beta_2 = 0$. What is the degree of freedom?

11. *(Regression Variance.)* Consider $y \sim N(X\beta, \sigma^2 I_n)$. In the class, we showed that $\sum_{i=1}^n \hat{e}_i^2/\sigma^2 \sim \chi^2_{n-p}$.

(a) An unbiased estimator for σ^2 is $\hat{\sigma}^2 = \frac{1}{n-p}\sum_{i=1}^n \hat{e}_i^2$. Find the asymptotic distribution of $\sqrt{n-p}(\hat{\sigma}^2 - \sigma^2)$.

(b) Find a transformation g, such that $\sqrt{n-p}(g(\hat{\sigma}^2) - g(\sigma^2))$ converges to a distribution that does not depend on σ^2.

(c) Can you do the above two questions for the intercept model?

12. *(Experiment Design.)* Consider the simple linear model $y_i \sim N(\beta_0 + \beta_1 x_i, \sigma^2)$ for $i = 1, 2, 3, 4$. Before you collect the four data points, you have freedom to choose x_1, x_2, x_3, x_4 and then observe y_1, y_2, y_3, y_4. Suppose you can only choose x_1, x_2, x_3, x_4 in the range $[-1, 1]$. How would you choose x_1, x_2, x_3, x_4 so that the accuracy of the MLE/LSE is the best?

13. *(Pythagorean Identity.)* Review and prove the following lists of Pythagorean identities:

(a) For any random variable $\hat{\theta}$ and any number θ, $\mathbb{E}(\hat{\theta} - \theta)^2 = \mathsf{Var}(\hat{\theta}) + (\mathbb{E}\hat{\theta} - \theta)^2$.

(b) For any real numbers $x_1, ..., x_n, \theta$, $\sum_{i=1}^n (x_i - \theta)^2 = \sum_{i=1}^n (x_i - \bar{x})^2 + n(\bar{x} - \theta)^2$.

(c) For any vectors $X_1, ..., X_n, \theta$, $\sum_{i=1}^n \|X_i - \theta\|^2 = \sum_{i=1}^n \|X_i - \bar{X}\|^2 + n\|\bar{X} - \theta\|^2$.

(d) For any $X \in \mathbb{R}^{n \times p}$, $y \in \mathbb{R}^{n \times 1}$ and $\mathbb{R}^{p \times 1}$, $\|y - X\beta\|^2 = \|y - X\hat{\beta}\|^2 + \|X\hat{\beta} - X\beta\|^2$, where $\hat{\beta} = (X^T X)^{-1} X^T y$.

(e) TSS = RSS + MSS.

14. Teach linear regression to your friend who does not know linear regression.