

# **ESMA 6787: Exam 2**

Due on December 14, 2025

*Israel Almodovar*

**Alejandro Ouslan**

## Problem 1: Acceptance of Syllabus

I have read the syllabus, understand its contents, and have no questions.

## Problem 2: Definitions

- (a) **Sample Space:** The sample space is the set of all possible outcomes of an experiment. It is to say everything that could happen in the experiment
- (b) **Kolmogorov Axioms of Probability:** It is a set of rules that define the probability of an event. Those rules can be summarized by the sum of the probability of all the events is 1, the probability of a given event should be non negative and if the events are disjoint then the probability of the union of the events is the sum of the probability of the events
- (c) **Exponential family:** A family of distributions is said to be an exponential family if the pdf can be written in the form
- (d) **Convergence in distribution:** Convergence of a distributions shows that as the size of given set of random outcome it the shape or distribution of that set will converge to a given distribution
- (e) **Convergence in Probability:** Convergence in Probability tells that as your sample gets larger the probability of a outcome will get close or more accurate compared to the original probability distribution
- (f) **Almost sure convergence (or convergence with probability 1):** Almost sure convergence tells if you were to follow an outcome repeated times you will almost always see it converge to something.
- (g) **Weak law of large numbers:** It tells that as you sample size increases your estimator will be closer to true estimator
- (h) **Strong law of large numbers:** If had infinite samples that your estimator will converge to the true estimator.
- (i) **Characteristics functions:** They are representations of the distributions.

## Problem 3:

Show that each of the following families of distributions is an exponential family,

- (a) The family of Bernoulli distribution with a unknown value of the parameter  $p$ .

$$f(x; p) = p^x (1-p)^{1-x}, \quad x \in \{0, 1\}$$

$$f(x; p) = \exp(x \log(p) + (1-x) \log(1-p))$$

Hence, it is in the form of an exponential family with natural parameter  $\theta = \log\left(\frac{p}{1-p}\right)$ , sufficient statistic  $T(x) = x$ , and log-partition function  $A(\theta) = \log(1 + \exp(\theta))$ .

- (b) The family of Poisson distributions with an unknown mean  $\lambda$

$$f(x; \lambda) = \frac{\lambda^x e^{-\lambda}}{x!}, \quad x \in \{0, 1, 2, \dots\}$$

$$f(x; \lambda) = \exp(x \log(\lambda) - \lambda - \log(x!))$$

Hence, it is in the form of an exponential family with natural parameter  $\theta = \log(\lambda)$ , sufficient statistic  $T(x) = x$ , and log-partition function  $A(\theta) = e^\theta$ .

- (c) The family of negative binomial distributions for which the value of  $r$  is known and the value of  $p$  is unknown.

$$f(x; p) = \binom{x+r-1}{x} p^x (1-p)^r, \quad x = 0, 1, 2, \dots$$

$$f(x; p) = \exp \left( x \log(p) + (r-1) \log(1-p) + \log \left( \binom{x+r-1}{x} \right) \right)$$

Thus, it is in the form of an exponential family with natural parameter  $\theta = \log \left( \frac{p}{1-p} \right)$ , sufficient statistic  $T(x) = x$ , and log-partition function  $A(\theta) = -r \log(1 - \exp(\theta))$ .

- (d) The family of normal distributions with an unknown mean and a known variance  $\sigma^2$ .

$$f(x; \mu) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left( -\frac{(x-\mu)^2}{2\sigma^2} \right)$$

$$f(x; \mu) = \exp \left( -\frac{1}{2\sigma^2} (x^2 - 2\mu x + \mu^2) \right)$$

Hence, it is in the form of an exponential family with natural parameter  $\theta = \frac{\mu}{\sigma^2}$ , sufficient statistic  $T(x) = x$ , and log-partition function  $A(\theta) = \frac{\mu^2}{2\sigma^2}$ .

- (e) The family of normal distributions with an unknown variance and a known mean.

$$f(x; \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left( -\frac{(x-\mu)^2}{2\sigma^2} \right)$$

$$f(x; \sigma^2) = \exp \left( -\frac{1}{2\sigma^2} (x-\mu)^2 - \frac{1}{2} \log(2\pi\sigma^2) \right)$$

Hence, it is in the form of an exponential family with natural parameter  $\theta = -\frac{1}{2\sigma^2}$ , sufficient statistic  $T(x) = (x-\mu)^2$ , and log-partition function  $A(\theta) = -\frac{1}{2} \log(-2\pi\theta)$ .

- (f) The family of gamma distributions for which the value of  $\alpha$  is unknown and the value of  $\beta$  is known.

$$f(x; \alpha) = \frac{x^{\alpha-1} e^{-x/\beta}}{\Gamma(\alpha)\beta^\alpha}, \quad x > 0$$

$$f(x; \alpha) = \exp \left( (\alpha-1) \log(x) - \frac{x}{\beta} - \log(\Gamma(\alpha)) - \alpha \log(\beta) \right)$$

Hence, it is in the form of an exponential family with natural parameter  $\theta = -\frac{1}{\beta}$ , sufficient statistic  $T(x) = x$ , and log-partition function  $A(\theta) = -\alpha \log(-\theta\beta)$ .

- (g) The family of gamma distributions for which the value of  $\alpha$  is known and the value of  $\beta$  is unknown.

$$f(x; \beta) = \frac{x^{\alpha-1} e^{-x/\beta}}{\Gamma(\alpha)\beta^\alpha}, \quad x > 0$$

$$f(x; \beta) = \exp \left( (\alpha-1) \log(x) - \frac{x}{\beta} - \log(\Gamma(\alpha)) - \alpha \log(\beta) \right)$$

Hence, it is in the form of an exponential family with natural parameter  $\theta = -\frac{1}{\beta}$ , sufficient statistic  $T(x) = x$ , and log-partition function  $A(\theta) = \alpha \log(-\theta)$ .

- (h) The family of beta distributions for which the value of  $\alpha$  is unknown and the value of  $\beta$  is known.

$$f(x; \alpha) = \frac{x^{\alpha-1} (1-x)^{\beta-1}}{B(\alpha, \beta)}, \quad x \in [0, 1]$$

$$f(x; \alpha) = \exp ((\alpha-1) \log(x) + (\beta-1) \log(1-x) - \log(B(\alpha, \beta)))$$

Hence, it is in the form of an exponential family with natural parameter  $\theta_1 = \alpha - 1$  and  $\theta_2 = \beta - 1$ , sufficient statistic  $T(x) = (\log(x), \log(1-x))$ , and log-partition function  $A(\theta_1, \theta_2) = \log(B(\alpha, \beta))$ .

- (i) The family of beta distributions for which the value of  $\alpha$  is known and the value of  $\beta$  is unknown.

$$f(x; \beta) = \frac{x^{\alpha-1}(1-x)^{\beta-1}}{B(\alpha, \beta)}, \quad x \in [0, 1]$$

$$f(x; \beta) = \exp((\alpha - 1)\log(x) + (\beta - 1)\log(1-x) - \log(B(\alpha, \beta)))$$

Hence, it is in the form of an exponential family with natural parameter  $\theta_1 = \alpha - 1$  and  $\theta_2 = \beta - 1$ , sufficient statistic  $T(x) = (\log(x), \log(1-x))$ , and log-partition function  $A(\theta_1, \theta_2) = \log(B(\alpha, \beta))$ .

### Problem 4:

Let  $X$  be a random variable with a Student's  $t$  distribution with  $p$  degrees of freedom.

- (a) Derive the mean and variance of  $X$ .

$$f(x) = \frac{\Gamma(\frac{p+1}{2})}{\sqrt{p\pi}\Gamma(\frac{p}{2})} \left(1 + \frac{x^2}{p}\right)^{-\frac{p+1}{2}}$$

$$E[X] = \int_{-\infty}^{\infty} x \cdot f(x) dx$$

$$g(-x) = (-x) \cdot f(-x) = -x \cdot f(x) = -g(x)$$

$$E[X] = 0 \quad \text{for } p > 1$$

- (b) Show that  $X^2$  has an  $F_{1,p}$  distribution.

$$X = \frac{Z}{\sqrt{V/p}}$$

where  $Z \sim N(0, 1)$  and  $V \sim \chi^2(p)$  are independent.

$$X^2 = \frac{Z^2}{V/p}$$

Since the square of a standard normal variable is a Chi-squared variable with 1 degree of freedom,  $Z^2 \sim \chi^2(1)$ .

$$X^2 = \frac{Z^2/1}{V/p}$$

The definition of the  $F$ -distribution with  $d_1$  and  $d_2$  degrees of freedom is:

$$F_{d_1, d_2} = \frac{U_1/d_1}{U_2/d_2}$$

where  $U_1 \sim \chi^2(d_1)$  and  $U_2 \sim \chi^2(d_2)$  are independent.

Identifying  $U_1 = Z^2$ ,  $d_1 = 1$ ,  $U_2 = V$ , and  $d_2 = p$ , we get:

$$X^2 \sim F_{1,p}$$

- (c) Let  $f(x|p)$  denote the pdf of  $X$ . Show that

$$\lim_{p \rightarrow \infty} f(x|p) \rightarrow \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{x^2}{2}\right\}$$

at each value of  $x$ ,  $-\infty < x < \infty$ . This correctly suggests that as  $p \rightarrow \infty$ ,  $X$  converges in distributions to a  $N(0, 1)$  random variable (Hint: Use Stirling's Formula).

**Answer:**

$$f(x|p) = \frac{\Gamma\left(\frac{p+1}{2}\right)}{\sqrt{p\pi}\Gamma\left(\frac{p}{2}\right)} \left(1 + \frac{x^2}{p}\right)^{-\frac{p+1}{2}}$$

Since  $\lim_{n \rightarrow \infty} (1 + \frac{a}{n})^n = e^a$ :

$$\lim_{p \rightarrow \infty} \left(1 + \frac{x^2}{p}\right)^{-\frac{p+1}{2}} = \lim_{p \rightarrow \infty} \left[\left(1 + \frac{x^2}{p}\right)^p\right]^{-\frac{p+1}{2p}} = (e^{x^2})^{-1/2} = e^{-x^2/2}$$

Using Stirling's approximation  $\frac{\Gamma(n+k)}{\Gamma(n)} \sim n^k$  as  $n \rightarrow \infty$ : Let  $n = p/2$  and  $k = 1/2$ :

$$\frac{\Gamma\left(\frac{p+1}{2}\right)}{\Gamma\left(\frac{p}{2}\right)} \approx \left(\frac{p}{2}\right)^{1/2}$$

$$\lim_{p \rightarrow \infty} \frac{1}{\sqrt{p\pi}} \frac{\Gamma\left(\frac{p+1}{2}\right)}{\Gamma\left(\frac{p}{2}\right)} = \frac{1}{\sqrt{p\pi}} \sqrt{\frac{p}{2}} = \frac{1}{\sqrt{2\pi}}$$

$$\lim_{p \rightarrow \infty} f(x|p) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$$

- (d) Use the results of parts (a) and (b) to argue that, as  $p \rightarrow \infty$ ,  $X$  converges in distribution to a  $X_1^2$  random variable.  
(e) What might you conjecture about the distributional limit, as  $p \rightarrow \infty$ , of  $F_{p,q}$

**Optional Problem 1:**

Suppose that  $X$  has the log-normal distribution with parameters  $\mu$  and  $\sigma^2$ . Find the distribution of  $\frac{1}{X}$ . Let  $X$  follow a lognormal distribution, so:

$$X \sim \text{Lognormal}(\mu, \sigma^2) \quad \text{implies} \quad \log(X) \sim N(\mu, \sigma^2)$$

This means that:

$$X = \exp(Z) \quad \text{where} \quad Z \sim N(\mu, \sigma^2)$$

Now, we are asked to find the distribution of  $Y = \frac{1}{X}$ .We can write  $Y$  as:

$$Y = \frac{1}{X} = \frac{1}{\exp(Z)} = \exp(-Z)$$

Since  $Z \sim N(\mu, \sigma^2)$ , the random variable  $-Z$  follows a normal distribution with parameters  $-\mu$  and  $\sigma^2$ , i.e.,

$$-Z \sim N(-\mu, \sigma^2)$$

Thus,  $Y = \exp(-Z)$  follows a lognormal distribution, and we conclude:

$$Y \sim \text{Lognormal}(-\mu, \sigma^2)$$

Therefore, the distribution of  $Y = \frac{1}{X}$  is a lognormal distribution with parameters  $-\mu$  and  $\sigma^2$ :

$$Y \sim \text{Lognormal}(-\mu, \sigma^2)$$

**Optional Problem 2:**

Suppose  $\bar{X}$  is the mean of 100 observations from a population with mean  $\mu$  and variance  $\sigma^2 = 9$ . Find limits between which  $\bar{X} - \mu$  will lie with probability at least 90%. Use both Chebyshev's inequality and the Central Limit Theorem, and comment on each.