University of Michigan, Dept of Statistics

Stats 510, Instructor: Long Nguyen

Solution to Midterm 2

Please do not distribute.

1. [12 points] Let function F be defined as follows:

$$F(x) = \begin{cases} \frac{1}{2}x^2, & x \in [0,1), \\ 0, & x < 0, \\ 1, & x \geq 1 \end{cases}$$

(i) [4pts] Show that F is a valid CDF for a random variable X.

(ii) [4pts] Find $\mathbb{E}X$ and $Var(X)$.

(iii) [4pts] What is F^{-1} in this case? Find all $x \in \mathbb{R}$ for which $F^{-1}(F(x)) = x$ holds.

Solution:

(i) F is non-decreasing and $\lim_{x\to-\infty} F(x) = 0$ and $\lim_{x\to\infty} F(x) = 1$. Also, since F is continuous when $x \neq 1$ and $F(1) = 1 = F(1+)$, F is right-continuous. Thus, F is a valid CDF.

(ii) $f(x) = x$, $x \in [0,1)$, and there is a pointmass on $X = 1$, so$P(X = 1) = 1/2$, and otherwise 0, by taking derivative on F. Thus,

$$\mathbb{E}X = \int_0^1 xf(x)dx + 1 \cdot P(X=1) = \int_0^1 x^2 dx + \frac{1}{2} = \frac{5}{6},$$

$$\mathbb{E}X^2 = \int_0^1 x^2 f(x)dx + 1 \cdot P(X=1) = \int_0^1 x^3 dx + \frac{1}{2} = \frac{3}{4},$$

$$Var(X) = \mathbb{E}X^2 - (\mathbb{E}X)^2 = \frac{3}{4} - \frac{25}{36} = \frac{1}{18}.$$

(iii) Recall the definition:

$$F^{-1}(p) = \inf\{x \in \mathbb{R} : F(x) \geq p\}.$$

For $0 \leq p < 1/2$,

$$F^{-1}(p) = \inf\{x \in \mathbb{R} : F(x) \geq p\} = \sqrt{2p}.$$

For $1/2 \leq p \leq 1$,

$$F^{-1}(p) = \inf\{x \in \mathbb{R} : F(x) \geq p\} = 1.$$

Thus,

$$F^{-1}(p) = \begin{cases} \sqrt{2p}, & 0 \leq p < 1/2, \\ 1, & 1/2 \leq p \leq 1. \end{cases}$$

If $0 \leq x < 1$, since $F(x) = \frac{1}{2}x^2 < \frac{1}{2}$,

$$F^{-1}(F(x)) = \sqrt{2F(x)} = \sqrt{x^2} = x.$$

If $x \geq 1$, since $F(x) = 1$,

$$F^{-1}(F(x)) = F^{-1}(1) = 1.$$

Therefore, $F^{-1}(F(x)) = x$ holds for $x \in [0,1]$.

Common Mistakes:

- For (i), since $F(x)$ is not continuous at $x = 1$, right continuity at $x = 1$ should be argued.
- For (ii), most of the students didn't take account the point mass of probability $1/2$ at $x = 1$. (2 pts given for the correct calculation except the point mass $\mathbb{E}X = \frac{1}{3}, Var(X) = \frac{5}{36}$)
- For (iii), many people missed to include 1.

2. [12 points] A function f on $\mathcal{R}$ is called symmetric about a if $f(a + x) = f(a - x)$ for all $x > 0$.

   (i) [4pts] Show that if the pdf of a random variable X is symmetric around a, and $\mathbb{E}X$ exists, then $\mathbb{E}X = a$.

   (ii) [4pts] Show that the pdf for an exponential distribution with parameter $\lambda > 0$: $f(x|\lambda) = \frac{1}{\lambda}e^{-x/\lambda}$, $x > 0$ and $f = 0$ otherwise, is not symmetric.

   (iii) [4pts] Let b be the median of an exponential random variable, i.e., b that satisfies $P(X > b) = P(X < b) = 1/2$. Show that $\mathbb{E}X > b$.

Solution:

(i)

$$\begin{aligned}
\mathbb{E}X &= \int xf(x)dx \\
&= a + \int (x-a)f(x)dx \\
&= a + \int_{-\infty}^{a} (x-a)f(x)dx + \int_{a}^{\infty} (x-a)f(x)dx \\
&= a - \int_{0}^{\infty} yf(a-y)dy + \int_{0}^{\infty} yf(a+y)dy \\
&= a + \int_{0}^{\infty} y(f(a+y) - f(a-y))dy \\
&= a.
\end{aligned}$$

(ii) If f is symmetric, it is symmetric around $\mathbb{E}X = \lambda$. Then it must holds that:

$$f(\lambda - x) = f(\lambda + x),$$

for all $x > 0$.

However, it is obvious that the equation doesn't hold for $x > \lambda$, since $f(\lambda - x) = 0$ and $f(\lambda + x) > 0$.

(iii)

$$P(X > b) = \int_{b}^{\infty} \frac{1}{\lambda}e^{-x/\lambda}dx = e^{-b/\lambda} = \frac{1}{2},$$

so the median is $b = \log 2 \cdot \lambda$.

$$\mathbb{E}X = \lambda > \log 2 \cdot \lambda = b.$$

Common Mistakes:

-

3. [12 points] Construct a sequence of non-identically distributed random variables $(X_n)_{n=1}^{\infty}$ that converges to a random variable Y in distribution in the following settings. (Specify both X_n and Y, and briefly justify your answer.)

   (i) [4pts] For each n, X_n is a discrete random variable that takes two possible values, and so is Y.

   (ii) [4pts] For each n, X_n is an exponential random variable, and so is Y.

   (iii) [4pts] For each each n, X_n is a discrete random variable that takes n possible distinct values.

   Solution:

   (i) **Two–point discrete.** Let $X_n \sim \text{Bernoulli}(p_n)$ with $p_n \to p \in [0,1]$, and let $Y \sim \text{Bernoulli}(p)$. Then the moment generating function of X_n converges to the moment generating function of Y:

   $$M_{X_n}(t) = 1 - p_n + p_n e^t \to 1 - p + pe^t = M_Y(t),$$

   hence X_n converges to Y in distribution.

   (ii) **Exponential.** Let $X_n \sim \text{Exp}(\lambda_n)$ with $\lambda_n \to \lambda > 0$, and let $Y \sim \text{Exp}(\lambda)$. For $x \geq 0$,

   $$F_{X_n}(x) = 1 - e^{-\lambda_n x} \longrightarrow 1 - e^{-\lambda x} = F_Y(x),$$

   hence X_n converges to Y in distribution.

   (iii) **n-point discrete.** Let X_n be uniform on the grid $\{1/n, 2/n, \ldots, n/n\}$:

   $$\mathbb{P}\left(X_n = \frac{k}{n}\right) = \frac{1}{n}, \qquad k = 1, \ldots, n.$$

   Let $Y \sim \text{Unif}(0,1)$. For $x \in (0,1)$,

   $$F_{X_n}(x) = \frac{\lfloor nx \rfloor}{n} \longrightarrow x = F_Y(x),$$

   where $\lfloor nx \rfloor = \max\{k \in \mathbb{Z} | k \leq nx\}$, hence X_n converges to Y in distribution.

   Remark:

   - The answer just gives one possible way for construction. Any valid answer works.
   - For justifying your answer, you can prove either the mgf converges or the cdf converges.

   Common Mistakes:

   - For (i), $X_n \sim \text{Bernoulli}(\frac{1}{n})$ converges to $Y = 1$, but Y only contains one possible value. (2 pts deducted)
   - For (iii), $X_n = Unif\{1, 2, \ldots, n\}$ does not converge to any distribution since $\mathbb{P}(X_n \leq a) \to 0$ for any $a > 0$ (2 pts deducted)
   - For (i) and (ii), if all the X_n have identical distribution, then no points are given.
   - For (iii), $Binomial(n,p)$ have $n+1$ possible values: $\{0, 1, \ldots, n\}$. No points is deducted for this small mistake but should keep in mind.

4. [14 points]Let $n > 0$ and function f be given as follows, for $x \in (-\infty, \infty)$:

   $$f(x) = \frac{n}{4} e^{-n|x-1|} + \frac{\sqrt{n}}{2\sqrt{2\pi}} e^{-\frac{n}{2}x^2}.$$

(i) [4pt] Show that f is a valid pdf for a random variable X.

(ii) [4pt] Show that the moment generating function exists for a neighborhood of 0 (that is, $M_X(t) < \infty$ for $t \in (-t_0, t_0)$ for some $t_0 > 0$).

(iii) [4pt] For $t \in (-t_0, t_0)$ from part (ii), find the integral expression for $\frac{d}{dt}M_X(t)$, and evaluate it at $t = 0$. Justify your operations, e.g., when interchanging the order of integrals and derivatives.

(iv) [2pt] Let X_n denote the random variable with the pdf f (to signify the fact that f depends on n). What does X_n converge in distribution to as n tends to infinity?

Solution:

(i) **Show that f is a valid pdf.**

We check two conditions: nonnegativity and normalization.

Clearly $f(x) \geq 0$ for all x because both terms are positive.

We compute

$$\int_{-\infty}^{\infty} f(x)\,dx = \int_{-\infty}^{\infty} \frac{n}{4} e^{-n|x-1|}\,dx + \int_{-\infty}^{\infty} \frac{\sqrt{n}}{2\sqrt{2\pi}} e^{-\frac{n}{2}x^2}\,dx.$$

For the first integral, split at $x = 1$:

$$\begin{aligned}\int_{-\infty}^{\infty} \frac{n}{4} e^{-n|x-1|}\,dx &= \frac{n}{4}\left(\int_{-\infty}^{1} e^{-n(1-x)}dx + \int_{1}^{\infty} e^{-n(x-1)}dx\right)\\ &= \frac{n}{4}\left(\frac{1}{n} + \frac{1}{n}\right) = \frac{1}{2}.\end{aligned}$$

For the second integral, recall $\int_{-\infty}^{\infty} e^{-\frac{n}{2}x^2}\,dx = \sqrt{\frac{2\pi}{n}}$ (using the density of normal distribution). Hence

$$\int_{-\infty}^{\infty} \frac{\sqrt{n}}{2\sqrt{2\pi}} e^{-\frac{n}{2}x^2}\,dx = \frac{\sqrt{n}}{2\sqrt{2\pi}} \cdot \sqrt{\frac{2\pi}{n}} = \frac{1}{2}.$$

Adding the two results gives $\int f(x)\,dx = 1$. Therefore f is a valid pdf.

(ii) **Show that the mgf exists in a neighborhood of 0.**

The moment generating function is

$$M_X(t) = \int_{-\infty}^{\infty} e^{tx} f(x)\,dx = I_1(t) + I_2(t),$$

where

$$I_1(t) = \int_{-\infty}^{\infty} e^{tx}\frac{n}{4}e^{-n|x-1|}\,dx, \qquad I_2(t) = \int_{-\infty}^{\infty} e^{tx}\frac{\sqrt{n}}{2\sqrt{2\pi}}e^{-\frac{n}{2}x^2}\,dx.$$

First term $I_1(t)$: Split at $x = 1$.

$$\begin{aligned}I_1(t) &= \frac{n}{4}\left(\int_{-\infty}^{1} e^{tx}e^{-n(1-x)}\,dx + \int_{1}^{\infty} e^{tx}e^{-n(x-1)}\,dx\right)\\ &= \frac{n}{4}\left(e^{-n}\int_{-\infty}^{1} e^{(t+n)x}\,dx + e^{n}\int_{1}^{\infty} e^{(t-n)x}\,dx\right).\end{aligned}$$

Each integral converges when $|t| < n$. Compute explicitly:

$$\int_{-\infty}^{1} e^{(t+n)x}\,dx = \frac{e^{(t+n)}}{t+n},$$
$$\int_{1}^{\infty} e^{(t-n)x}\,dx = \frac{e^{(t-n)}}{n-t}.$$

Substitute:

$$I_1(t) = \frac{n}{4}\left(e^{-n}\frac{e^{t+n}}{t+n} + e^{n}\frac{e^{t-n}}{n-t}\right) = \frac{ne^t}{4}\left(\frac{1}{t+n} + \frac{1}{n-t}\right) = \frac{n^2e^t}{2(n^2-t^2)}.$$

Thus $I_1(t)$ is finite for $|t| < n$.

Second term $I_2(t)$:

$$I_2(t) = \frac{\sqrt{n}}{2\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{tx}e^{-\frac{n}{2}x^2}\,dx.$$

Complete the square:

$$tx - \frac{n}{2}x^2 = -\frac{n}{2}\left(x^2 - \frac{2t}{n}x\right) = -\frac{n}{2}\left(x - \frac{t}{n}\right)^2 + \frac{t^2}{2n}.$$

Hence

$$\begin{aligned} I_2(t) &= \frac{\sqrt{n}}{2\sqrt{2\pi}}e^{\frac{t^2}{2n}} \int_{-\infty}^{\infty} e^{-\frac{n}{2}\left(x-\frac{t}{n}\right)^2}\,dx = \frac{\sqrt{n}}{2\sqrt{2\pi}}e^{\frac{t^2}{2n}}\sqrt{\frac{2\pi}{n}} \\ &= \frac{1}{2}e^{\frac{t^2}{2n}}. \end{aligned}$$

Thus the full mgf is

$$M_X(t) = I_1(t) + I_2(t) = \frac{n^2e^t}{2(n^2-t^2)} + \frac{1}{2}e^{\frac{t^2}{2n}},$$

which is finite for all $|t| < n$. Therefore, the mgf exists in a neighborhood of 0.

(iii) **Find an integral expression for $M_X'(t)$ and evaluate it at $t = 0$.**

From the definition

$$M_X(t) = \int_{\mathbb{R}} e^{tx} f(x)\,dx = \int_{\mathbb{R}} e^{tx}\left(\frac{n}{4}e^{-n|x-1|} + \frac{\sqrt{n}}{2\sqrt{2\pi}}e^{-\frac{n}{2}x^2}\right)dx =: I_1(t) + I_2(t).$$

Justification to differentiate under the integral. Fix $t_0 \in (0, n)$. For $|t| \le t_0$,

$$|x|\,e^{tx}\frac{n}{4}e^{-n|x-1|} \;\le\; C_1\,(1+|x|)\,e^{-(n-t_0)|x|},$$

and

$$|x|\,e^{tx}\frac{\sqrt{n}}{2\sqrt{2\pi}}e^{-\frac{n}{2}x^2} \;\le\; C_2\,(1+|x|)\,e^{-\frac{n}{4}x^2},$$

for suitable constants C_1, C_2 (use $e^{tx} \le e^{t_0|x|}$ and split at $x = 1$ for the first term, and complete the square for the second). Hence by the Dominated Convergence Theorem (Leibniz rule),

$$M_X'(t) = \int_{\mathbb{R}} x\,e^{tx}\,f(x)\,dx \qquad (|t| < t_0 < n).$$

Evaluation at $t = 0$. Then

$$M_X'(0) = \int_{\mathbb{R}} x\, f(x)\, dx = \frac{n}{4} \int_{\mathbb{R}} x\, e^{-n|x-1|} dx + \frac{\sqrt{n}}{2\sqrt{2\pi}} \int_{\mathbb{R}} x\, e^{-\frac{n}{2}x^2} dx.$$

The second integral is 0 (odd integrand). For the first, substitute $y = x - 1$:

$$\int_{\mathbb{R}} x\, e^{-n|x-1|} dx = \int_{\mathbb{R}} (y+1)\, e^{-n|y|} dy = \underbrace{\int_{\mathbb{R}} y\, e^{-n|y|} dy}_{=0} + \int_{\mathbb{R}} e^{-n|y|} dy = \frac{2}{n}.$$

Therefore

$$M_X'(0) = \frac{n}{4} \cdot \frac{2}{n} + 0 = \frac{1}{2}.$$

(Cross-check via explicit M_X.) From part (ii) we also had, for $|t| < n$,

$$M_X(t) = \frac{n^2 e^t}{2(n^2 - t^2)} + \frac{1}{2}\, e^{\frac{t^2}{2n}}.$$

Differentiating this closed form,

$$M_X'(t) = \frac{n^2}{2} e^t \left(\frac{1}{n^2 - t^2} + \frac{2t}{(n^2 - t^2)^2} \right) + \frac{1}{2} e^{\frac{t^2}{2n}} \frac{t}{n},$$

so $M_X'(0) = \frac{n^2}{2} \cdot \frac{1}{n^2} + \frac{1}{2} \cdot 0 = \frac{1}{2}$, agreeing with the integral computation.

(iv) **Limit distribution of X_n.**
From part (ii) we obtained, for $|t| < n$,

$$M_{X_n}(t) = \frac{n^2 e^t}{2\,(n^2 - t^2)} + \frac{1}{2} \exp\left(\frac{t^2}{2n}\right).$$

Fix any $t \in \mathbb{R}$. For all sufficiently large n we have $|t| < n$, so

$$\lim_{n\to\infty} M_{X_n}(t) = \frac{1}{2} e^t \cdot \lim_{n\to\infty} \frac{n^2}{n^2 - t^2} + \frac{1}{2} \cdot \lim_{n\to\infty} \exp\left(\frac{t^2}{2n}\right) = \frac{1}{2} e^t + \frac{1}{2} =: M(t).$$

The pointwise limit $M(t) = \frac{1}{2}(1 + e^t)$ is finite for all t, satisfies $M(0) = 1$, and is the mgf of a Bernoulli(1/2) random variable Y (indeed, $M_Y(t) = \mathbb{E}[e^{tY}] = \frac{1}{2}e^0 + \frac{1}{2}e^t$). By the mgf convergence theorem (Curtiss' theorem), since $M_{X_n}(t) \to M(t)$ on a neighborhood of 0 and M is an mgf, we conclude that X_n converges to Y in distribution.

Remark:

- For (i) and (ii), if you observe f is a mixture of Laplace and Normal distribution, you don't need to to do the detailed computing. you can just use the mgf of this two distributions to give the mgf of f.
- For (iii), you can either find the moment generating function using the above way (i.e., changing the operations of integral and differential), or just directly compute the differential from (ii).

Common Mistakes:

-