Bootcamp Exercises

Alex Dombowsky

June 2021

Exercise 1: the Normal Distribution and Moments

Let $X \sim \mathcal{N}(0,1)$.

(a) Give the pdf for the distribution of X^2.

Using change-of-variables, this is a χ^2_1 (or a Gamma$(1/2, 1/2)$) distribution.

(b) Give an expression for $\mathbb{E}[X^k]$ for all odd k.

$$\mathbb{E}[X^k] = \int_{-\infty}^{\infty} (2\pi)^{-1/2} \underbrace{x^k}_{\text{Odd}} \underbrace{\exp\left\{-\frac{x^2}{2}\right\}}_{\text{Even}} dx = 0$$

since we are integrating an odd function over a symmetric interval.

(c) Let $Y = X^2$. Calculate Cov(X, X^2).

$$\text{Cov}(X, X^2) = \underbrace{\mathbb{E}[X^3]}_{=0} - \mathbb{E}[X^2] \underbrace{\mathbb{E}[X]}_{=0} = 0.$$

Note however that X and X^2 are *not independent*.

(d) The *moment generating function* of X is defined as $M_X(t) = \mathbb{E}[e^{tX}]$ for all $t \in \mathbb{R}$. Using integrals, find the closed form expression for $M_X(t)$.

$$M_X(t) = \mathbb{E}[e^{tX}] = \int (2\pi)^{-1/2} \exp\left\{-\frac{x^2}{2} + tx\right\} dx$$

$$= e^{t^2/2} \underbrace{\int (2\pi)^{-1/2} \exp\left\{-\frac{1}{2}(x-t)^2\right\} dx}_{=1} = e^{t^2/2}.$$

(e) Let $Z \sim \mathcal{N}(\mu, \sigma^2)$. Using part (d), find $M_Z(t) = \mathbb{E}[e^{tZ}]$ for all t.

$$Z \overset{d}{=} \mu + \sigma X \implies M_Z(t) = e^{\mu t} e^{\sigma^2 t^2/2} = e^{\mu t + \sigma^2 t^2/2}.$$

(f) **Optional Challenge**: Assume $\sigma^2 = 1$, so $Z \sim \mathcal{N}(\mu, 1)$. The *characteristic function* of Z is defined as $\varphi_Z(t) = \mathbb{E}[e^{itZ}]$ where $i = \sqrt{-1}$. Fortunately, $\varphi_Z(t) = M_Z(it)$. Find an unbiased estimator for $(-1)^\mu$. *Hint*: you may find *Euler's Identity* to be helpful here, which is $e^{i\pi} = -1$.

Exercise 2: Classic 1-1 Transformations of Continuous Random Variables

Let $X_1 \sim \text{Gamma}(a, \xi)$ and $X_2 \sim \text{Gamma}(b, \xi)$ where $a, b, \xi > 0$ and $X_1 \perp\!\!\!\perp X_2$. For each problem give the pdf (*don't forget the support*) and the name of the distribution.

(a) What is the distribution of $W = \frac{X_1}{X_1+X_2}$?

Make it a one-to-one transformation with the variable $U = X_1$. This gives us inverse transformations

$$X_1 = U \text{ and } X_2 = U(W^{-1} - 1).$$

Note that the support of W is $(0, 1)$, while the support of U is $\mathbb{R}^+$. The Jacobian is

$$J = \begin{pmatrix} 0 & 1 \\ -u/w^2 & (w^{-1} - 1) \end{pmatrix}$$

with determiniant u/w^2. So, the joint distribution of (W, U) is

$$\begin{aligned} f_{W,U}(w, u) &= f_{X_1}(u) f_{X_2}(u(1/w - 1))u/w^2 \\ &= \frac{\xi^a}{\Gamma(a)} u^{a-1} e^{-\xi u} \frac{\xi^b}{\Gamma(b)} u^{b-1}(1/w - 1)^{b-1} e^{-\xi u(1/w-1)} u/w^2 \\ &= \frac{\xi^{a+b}}{\Gamma(a)\Gamma(b)} (1 - w)^{b-1} w^{-1-b} u^{a+b-1} e^{-(\xi/w)u}. \end{aligned}$$

So,

$$\begin{aligned} f_W(w) = \int_0^\infty f_{W,U}(w, u)du &= \frac{\xi^{a+b}}{\Gamma(a)\Gamma(b)} (1 - w)^{b-1} w^{-1-b} \underbrace{\int_0^\infty u^{a+b-1} e^{-(\xi/w)u} du}_{\text{Gamma}(a+b,\xi/w) \text{ kernel}} \\ &= \frac{\xi^{a+b}}{\Gamma(a)\Gamma(b)} (1 - w)^{b-1} w^{-1-b} \frac{\Gamma(a + b)}{\xi^{a+b}(1/w)^{a+b}} \\ &= \frac{\Gamma(a + b)}{\Gamma(a)\Gamma(b)} (1 - w)^{b-1} w^{a-1} \end{aligned}$$

for $0 < w < 1$. This is the pdf of a $\text{Beta}(a, b)$ distribution.

(b) Set $a = b = 1$. For $\lambda > 0$, what is the distribution of $Y = -\frac{1}{\lambda} \log W$?

Note that in this case, $W \sim \text{Unif}(0, 1)$. The inverse transformation is $W = e^{-\lambda Y}$, which has Jacobian (in absolute value) $\lambda e^{-\lambda y}$. This tells us that

$$f_Y(y) = \lambda \mathbf{1}(e^{-\lambda y} \in (0, 1))e^{-\lambda y} = \lambda e^{-\lambda y} \mathbf{1}(y > 0)$$

which is the pdf of a $\text{Exp}(1)$ distribution.

(c) What is the distribution of $Z = Y^{1/\alpha}$ for $\alpha > 0$?

The inverse transformation is $Y = Z^\alpha$ with Jacobian $\alpha z^{\alpha-1}$. This tells us that

$$f_Z(z) = \lambda e^{-\lambda z^\alpha} \mathbf{1}(z^\alpha > 0)\alpha z^{\alpha-1} = \lambda\alpha z^{\alpha-1} e^{-\lambda z^\alpha} \mathbf{1}(z > 0)$$

which is the pdf of a $\text{Weibull}(\alpha, \lambda)$ distribution.

(d) Let $U_1 \sim \chi^2_{\nu_1}$ and $U_2 \sim \chi^2_{\nu_2}$ with $U_1 \perp\!\!\!\perp U_2$. Show that $F = \frac{U_1/\nu_1}{U_2/\nu_2}$ has an $F(\nu_1, \nu_2)$ distribution.

Similar derivation to part (a) using the one-to-one transformation

$$F = \frac{U_1/\nu_1}{U_2/\nu_2} \text{ and } V = U_1.$$

The inverse transformations are

$$U_1 = V \text{ and } U_2 = \frac{V/\nu_1}{F/\nu_2}.$$

(e) Let X be a continuous random variable with support $\mathbb{R}$ with cdf $F_X(x)$. What is the distribution of $V = F_X(X)$?

Such conditions (continuity and support on the whole real line) mean that F is invertible. While you can use the change-of-variables approach, the easiest way is to use CDFs. That is, first note that $V \in (0, 1)$. Next, for $v \in (0, 1)$

$$F_V(v) = P(V \leq v) = P(F_X(X) \leq v) = P(X \leq F_X^{-1}(v)) = F_X(F_X^{-1}(v)) = v.$$

So, the pdf of V is

$$f_V(v) = \mathbf{1}(0 < v < 1)$$

which tells us that $V \sim \text{Unif}(0, 1)$.

(f) **Optional Challenge**: Prove or disprove–part (d) holds for discrete random variables.

Exercise 3: Fun with Exponentials

Let $X_1, \ldots, X_n \overset{iid}{\sim} \text{Exp}(\lambda)$.

(a) What is the distribution of $X_{(1)} = \min(X_1, \ldots, X_n)$?

$$\begin{aligned} P(X_{(1)} \leq x) &= 1 - P(X_{(1)} > x) = 1 - P(X_1 > x, \ldots, X_n > x) \\ &= 1 - \prod_{i=1}^{n}(1 - P(X_i \leq x)) = 1 - (e^{-\lambda})^n = 1 - e^{-n\lambda} \end{aligned}$$

which is the CDF of an $\text{Exp}(n\lambda)$ random variable.

(b) Derive the moment generating function for X_i.

$$\begin{aligned} M_{X_i}(t) = \mathbb{E}[e^{tX_i}] &= \int e^{tx} \lambda e^{-\lambda x} dx = \lambda \underbrace{\int e^{-(\lambda - t)x} dx}_{\text{Exp}(\lambda - t)} \\ &= \lambda \cdot \frac{1}{\lambda - t} = \frac{\lambda}{\lambda - t} \end{aligned}$$

provided $t < \lambda$. For $t \geq \lambda$, $M_X(t) = \infty$.

(c) What is the distribution of $Y = \sum_{i=1}^{n} X_i$?

$$M_Y(t) = \prod_{i=1}^{n} M_{X_i}(t) = \left(\frac{\lambda}{\lambda - t}\right)^n$$

which is the MGF of a $\text{Gamma}(n, \lambda)$ distribution.

(d) Let $W \sim \text{Poisson}(\mu)$, $X \sim \text{Exp}(\lambda)$, $X \perp\!\!\!\perp W$. Is $Z = X - W$ continuous, discrete, or neither?

Since X can take uncountably many values, W can take countably many values, and since $X \perp\!\!\!\perp W$, the support of Z is uncountable. Therefore, Z is continuous.

(e) Give an expression for $P(Z < z)$ for $z \in \mathbb{R}$.

For $z \in \mathbb{R}$,

$$
\begin{aligned}
P(Z < z) = P(Z \leq z) = \mathbb{E}[\mathbf{1}(X < W + z)] &= \sum_{w=0}^{\infty} \int_0^{\infty} \mathbf{1}(X < W + z) f_X(x) f_W(w) dx \\
&= \sum_{w=0}^{\infty} \left(\int_0^{w+z} \lambda e^{-\lambda x} dx \right) \frac{e^{-\mu} \mu^w}{w!} \\
&= \sum_{w=0}^{\infty} (1 - e^{-\lambda(w+z)}) \frac{e^{-\mu} \mu^w}{w!} = 1 - e^{-(\lambda z + \mu)} \sum_{w=0}^{\infty} \frac{(e^{-\lambda} \mu)^w}{w!} \\
&= 1 - e^{-\lambda z - \mu} e^{e^{-\lambda} \mu} = 1 - \exp\left\{ -\lambda z + (e^{-\lambda} - 1)\mu \right\}.
\end{aligned}
$$

(f) **Optional Challenge**: Let $n = 2$. Show that $T_1 = \min(X_1, X_2)$ and $T_2 = X_1 - X_2$ are independent.

Exercise 4: Heavy Tailed Distributions

Let $X_1, X_2 \overset{iid}{\sim} \mathcal{N}(0, 1)$, $W \sim \chi^2_\nu$, $W \perp\!\!\!\perp X_1, X_2$.

(a) Let $X_1, X_2 \overset{iid}{\sim} \mathcal{N}(0, 1)$. What is the distribution of $Y = \frac{X_1}{X_2}$? What is $\mathbb{E}[Y]$?

(b) What is the distribution of $T = \frac{X_1}{\sqrt{W/\nu}}$?

(c) Show that $T \overset{d}{=} Y$ if $\nu = 1$. That is, $\frac{X_1}{X_2} \overset{d}{=} \frac{X_1}{\sqrt{W}}$. Conceptually, why does this make sense?

Exercise 5: Large Sample Theory

(a) Let $X_1, \ldots, X_n \overset{iid}{\sim} \text{Exp}(1)$. Let $M_n = \max(X_1, \ldots, X_n)$. Find the limiting distribution (you can just state the CDF) of $M_n - \log(n)$.

Let $Z_n = M_n - \log(n)$. Then,

$$
\begin{aligned}
F_{Z_n}(z) = P(Z_n \leq z) = P(M_n \leq z + \log(n)) &= \prod_{i=1}^{n} P(X_i \leq z + \log(n)) = \left(1 - e^{-(z + \log(n))}\right)^n \\
&= \left(1 - \frac{e^{-z}}{n}\right)^n \to \exp\left\{-e^{-z}\right\}
\end{aligned}
$$

as $n \to \infty$.

(b) Let $X_1, \ldots, X_n$ be *iid* continuous random variables with pdf f and $\mathbb{E}[X_i] = \mu < \infty$. Let g be another pdf (with the same support as f). Show that

$$
W_n = \frac{\sum_{i=1}^n X_i}{\sum_{i=1}^n g(X_i)/f(X_i)} \xrightarrow{p} \mu
$$

as $n \to \infty$.

By the Weak Law of Large numbers,

$$\frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow{p} \mu \implies \frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow{d} \mu.$$

Similarly, the sample of random variables $g(X_i)/f(X_i)$ are *iid*, with expectation

$$\mathbb{E}[g(X_i)/f(X_i)] = \int \frac{g(x)}{f(x)} f(x)dx = \int g(x)dx = 1 < \infty.$$

So, again by the Weak Law,

$$\frac{1}{n}\sum_{i=1}^{n} g(X_i)/f(X_i) \xrightarrow{p} 1.$$

By Slutsky's theorem,

$$W_n = \frac{\sum_{i=1}^n X_i}{\sum_{i=1}^n g(X_i)/f(X_i)} = \frac{(1/n)\sum_{i=1}^n X_i}{(1/n)\sum_{i=1}^n g(X_i)/f(X_i)} \xrightarrow{p} \mu/1 = \mu.$$

(c) Let $f(x|\theta_1, \theta_2) = \frac{1}{\theta_2}\exp\left(-\frac{x-\theta_1}{\theta_2}\right), x \geq \theta_1, \theta_1 \in \mathbb{R}, \theta_2 > 0$. Find the maximum likelihood estimators for θ_1, θ_2

$\hat{\theta_1} = \min_i x_i, \hat{\theta_1} = \overline{x} - \min_i x_i$

(d) **Optional Challenge**: Let $X_1, \ldots, X_n \overset{iid}{\sim} \text{Cauchy}(0,1)$. What is the distribution of $Z_n = \frac{1}{n}\sum_{i=1}^n X_i$? *Hint: this requires the CF of the standard Cauchy distribution, which is* $\varphi(t) = \exp(-|t|)$. *This is a nice counterexample for when the* L_1 *requirement of the Central Limit Theorem does not hold.*

Exercise 6: True or False Questions

For each statement, provide a short answer or proof to why it is true or false.

1. For a random variable X, the moment generating function $M_X(t)$ exists (ie, is finite) for all $t \in \mathbb{R}$.

   False. While this is true for Gaussian random variables, the MGF may only be finite in an interval (such as the Gamma distribution) or only at a point (the Cauchy distribution MGF is finite only when $t = 0$). The characteristic function is finite (for any distribution) for all $t \in \mathbb{R}$, which you can show using basic complex analysis. You'll learn more about characteristic functions in STA 711.

2. Let $X \sim \mathcal{N}(0,1)$. Then $Y = \mathbf{1}(X > 0) \sim \text{Bernoulli}(1/2)$.

   True.

3. If X and Y are two random variables such that $X \sim N(\mu_x, \sigma_x^2), Y \sim N(\mu_y, \sigma_y^2)$, the pair (X, Y) has a joint normal distribution.

   False: Counter example is $X \sim N(0,1), Y = X if |X| > c, Y + -X$ if $|X| < c, c > 0$

4. If X and Y are two independent random variables such that $X \sim N(\mu_x, \sigma_x^2), Y \sim N(\mu_y, \sigma_y^2)$, then the pair (X, Y) has a bivariate normal distribution.

   True.

5. If X and Y are two normally distributed random variables that are uncorrelated, then they are independent

   False: let $Y = WX$ where W is a discrete random variable that takes on the values $\{-1, 1\}$ with equal probability.

6. If X and Y are jointly normal random variables that are uncorrelated, then they are independent. True.

7. Let X be a discrete random variable and $g : \mathbb{R} \to \mathbb{R}$ be some (well-defined) function. Then $Y = g(X)$ is discrete.

   True.

8. Let X be a continuous random variable and $g : \mathbb{R} \to \mathbb{R}$ be some function. Then $Y = g(X)$ is continuous.

   False.