Lab 3

Expectations, Integral Transforms, & Generating Functions

Abstract: We have introduced expectations in lecture and its associated ideas, such as characteristic / moment generating functions. To help motivate and better understand these concepts, we will start with a more general framework regarding integral transformations and generating functions. This lab is intended to also help encourage comfort when working with sums and series expansions.

Key words: expectation, probability generating function, moment generating function, characteristic function, integral transforms, Fourier transform, Laplace transform

3.1. Integral transforms (optional)

Definition 3.1.1. An *integral transform T* transforms $f \xrightarrow{T} f^*$ such that

$$f^*(u) = \int_{t_1}^{t_2} K(t,u)f(t)dt$$

where $K(t,u)$ is known as the *kernel function.* We say that there exists an *inverse kernel* $K^{-1}(u,t)$ if $K^{-1}(u,t)$ satisfies

$$f(u) = \int_{u_1}^{u_2} K^{-1}(u,t)f^*(u)du$$

Mathematical notation aside, the motivation behind integral transforms is easy to understand. There are many classes of problems that are difficult to solve—or at least quite unwieldy algebraically—in their original representations. An integral transform "maps" an equation from its original "domain" into another domain. Manipulating and solving the equation in the target domain can be much easier than manipulation and solution in the original domain. The solution is then mapped back to the original domain with the inverse of the integral transform.

The table below presents a few integral transforms and their respective (inverse) kernels.

Transform	Discrete/Continuous	t_1	t_2	K	K^{-1}
Two-sided Laplace transform	Continuous	$-\infty$	∞	e^{-ut}	$e^{ut}/(2\pi i)$
Fourier transform	Continuous	$-\infty$	∞	$e^{-2\pi iut}$	$e^{2\pi iut}$
Z transform	Discrete	$-\infty$	∞	u^{-t}	$u^{t-1}/(2\pi i)$
Binomial transform	Discrete	0	∞	$(-1)^{u-t}\binom{u}{t}$	$\binom{t}{u}$

Example 3.1.1. (Moments and Central Moments) For random variable X with finite moments, denote $\mu_n = \mathbb{E}(X-\mu)^n$ and $\mu'_n = \mathbb{E}X^n$, where $\mu = \mu_1 = \mathbb{E}X$. The binomial theorem instantly gives us

$$\mu_n = \sum_{k=0}^{n}(-1)^{n-k}\binom{n}{k}\mu'_k\mu^{n-k}$$

If we let $f^*(n) = \mu_n/\mu^n$ and $f(k) = \mu'_k/\mu^k$, we may write the above expression as

$$f^*(n) = \sum_{k=0}^{\infty}(-1)^{n-k}\binom{n}{k}f(k)$$

Notice that f^* is the binomial transform of f. We can apply the inverse to get

$$f(k) = \sum_{n=0}^{k}\binom{k}{n}f^*(n)$$

Or in other words,

$$\mu'_k = \sum_{n=0}^{k}\binom{k}{n}\mu_n\mu^{k-n}$$

Example 3.1.2. (PGF) Let X take values in $\mathbb{N}_0 = \{0,1,2,\cdots\}$ with PMF $p_X(x)$. Define $g_X(t)$ as the Z transform of $p_X(x)$; that is,

$$g_X(t) = \sum_{x=-\infty}^{\infty} t^{-x}p_X(-x) = \sum_{x=0}^{\infty} t^x p_X(x) = \mathbb{E}[t^X]$$

In fact, $g_X(t)$ is known as the *probability generating function* of X. We can invert to get

$$p_X(-x) = \int_{-\infty}^{\infty}\frac{t^{x-1}}{2\pi i}g_X(t)dt \implies p_X(x) = \int_{-\infty}^{\infty}\frac{1}{2\pi i t^{x+1}}g_X(t)dt$$

Example 3.1.3. (MGF) For random variable X, either continuous or discrete with pdf/cdf $f_X(x)$, let $M_X(t)$ be the two-sided Laplace transform of $f_X(-x)$. Then,

$$M_X(t) = \int_{-\infty}^{\infty} e^{-tx}f_X(-x)dx = \int_{-\infty}^{\infty} e^{tx}f_X(x)dx = \mathbb{E}[e^{tX}]$$

This is exactly the *moment generating function* for X. Applying the inverse integral transform, we get

$$f_X(-x) = \int_{-i\infty}^{i\infty} \frac{e^{tx}}{2\pi i} M_X(t)dt \implies f_X(x) = \int_{-i\infty}^{i\infty} \frac{e^{-tx}}{2\pi i} M_X(t)dt$$

where $\int_{-i\infty}^{i\infty}$ is interpreted as integrating along the imaginary axis.

Example 3.1.4. (CF) For random variable X, either continuous or discrete with pdf/cdf $f_X(x)$, let $\phi_X(t)$ be the Fourier transform of $2\pi f_X(-2\pi x)$. Then,

$$\phi_X(t) = \int_{-\infty}^{\infty} e^{-2\pi itx} 2\pi f_X(-2\pi x)dx = \int_{-\infty}^{\infty} e^{itx} f_X(x)dx = \mathbb{E}[e^{itX}]$$

This is exactly the *characteristic function* for X. Applying the inverse integral transform, we get

$$2\pi f_X(-2\pi x) = \int_{\infty}^{\infty} e^{2\pi itx}\phi_X(t)dt \implies f_X(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-itx}\phi_X(t)dt$$

The four previous examples demonstrate each of the transforms listed in the table, and show how if we have one form, we can always get the other. Integral transforms do not necessarily have to be invertible, but if they are, we are endowed with a ton of nice properties, but most relevant to us, they imply

Theorem 3.1.1. Suppose X, Y are random variables. Then

$$X \overset{\mathcal{D}}{=} Y \iff \phi_X(t) = \phi_Y(t) \quad \text{for all } t \text{ in some positive length open interval } O \subseteq \mathbb{R}$$

We could replace the CF above with either MGF or PGF, provided that these quantities exist. This theorem follows directly from the fact that there exists a bijection between the space of distribution functions and the space of CF's (or MGF's/PGF's, provided they exist).

3.2. Generating functions

The name *generating functions* is exactly what it means; these functions generate something. But the way this "something" is generated is encoded in its coefficients. For example, we can say that e^x generates the sequence $\{\frac{1}{n!}\}_{n=0}^{\infty}$. In a similar manner, $g_X(t)$ generates the sequence $\{p_X(x)\}_{x=0}^{\infty}$, and $M_X(t)$ generates the sequence $\{\mu_k'/k!\}_{k=0}^{\infty}$. Probability generating functions will touched upon later in the course. MGF's look almost like CF's, and in fact, if MGF's exist, then we have the relation $M_X(it) = \phi_X(t)$. For convenience, we state some properties regarding MGF's. Extensions to CF's can be worked through the relationship stated above.

Theorem 3.2.1. Suppose X has MGF M_X. We have

- If $Y = a + bX$, then $M_Y(t) = e^{at}M_X(bt)$.
- If X and Y are independent and $Z = X + Y$, then $M_Z(t) = M_X(t)M_Y(t)$.
- $\frac{d^n}{dt^n}M_X(t)\bigg|_{t=0} = \mathbb{E}X^n$

Proof. Proofs are as follows:

- $M_Y(t) = \mathbb{E}e^{tY} = \mathbb{E}e^{at}e^{btX} = e^{ta}\mathbb{E}e^{(bt)X} = e^{at}M_X(bt)$
- $M_Z(t) = \mathbb{E}e^{tZ} = \mathbb{E}e^{tX}e^{tY} = \mathbb{E}e^{tX}\mathbb{E}e^{tX} = M_X(t)M_Y(t)$, where the third equality uses independence to factor the expectation.
- We have

$$\begin{aligned}\frac{d^n}{dt^n}M_X(t) &= \frac{d^n}{dt^n}\sum_{k=0}^{\infty}\frac{t^k\mu_k'}{k!}\\ &= \sum_{k=0}^{\infty}\frac{d^n}{dt^n}\frac{t^k\mu_k'}{k!}\\ &= \sum_{k=n}^{\infty}\frac{(k)_n t^{k-n}\mu_k'}{k!}\end{aligned}$$

where $(k)_n = (k)(k-1)\cdots(k-n+1)$ is the falling factorial. Note that, for $k > n$, we have $t^{k-n}\big|_{t=0} = 0$, hence evaluating the expression at $t = 0$ produces $\mu_n' = \mathbb{E}X^n$.

□

3.3. Some worked problems

Example 3.3.1. (Tail-sum formula) For a positive CRV X, prove that $\mathbb{E}X = \int_0^\infty \mathbb{P}(X > t)dt$.

Proof. This is a simple exercise with Fubini's theorem:

$$\begin{aligned}\int_0^\infty \mathbb{P}(X > t)dt &= \int_0^\infty\int_t^\infty f_X(x)dxdt\\ &= \int_0^\infty\int_0^x f_X(x)dtdx\\ &= \int_0^\infty xf_X(x)dx\\ &= \mathbb{E}X\end{aligned}$$

□

Now here is a rather clever application of the (discrete) tail-sum formula:

Example 3.3.2. A merchant travels to a far away land, hoping to bring back the spoils of his journeys. However, his caravan is not the most competent and hence loses things. If he were to acquire weight A in gold coins at his destination, he is expected to bring home $U \cdot A$, where $U \sim \text{Unif}(0, 1)$. If the merchant acquires one ton in gold coins at his destination for every trip, and will not stop until he has at least one ton in gold coins back home, what is the expected number of trips he will make?

Solution. Let $U_1, \cdots, U_n \sim \text{Unif}(0,1)$, and let $N = \min\{n : \sum_{i=1}^n U_i \geq 1\}$. The problem is equivalent to finding $\mathbb{E}N$. We have

$$\begin{aligned}
\mathbb{E}N &= \sum_{n=0}^{\infty} \mathbb{P}(N > n) \\
&= 1 + \sum_{n=1}^{\infty} \mathbb{P}(N > n) \\
&= 1 + \sum_{n=1}^{\infty} \mathbb{P}\left(\sum_{i=1}^{n} U_i \leq 1\right) \\
&= 1 + \sum_{n=1}^{\infty} \int_{\sum_{i=1}^n u_i \leq 1} du_1 \cdots du_n
\end{aligned}$$

Note that $V_n = \int_{\sum_{i=1}^n u_i \leq 1} du_1 \cdots du_n$ corresponds to the volume under the n-dimensional standard tetrahedron (also known as a simplex). We can compute

$$\begin{aligned}
V_n &= \int_{u_n \leq 1} \left(\int_{\sum_{i=1}^{n-1} u_i \leq 1 - u_n} du_1 \cdots du_{n-1} \right) du_n \\
&= \int_{u_n \leq 1} \left(\int_{\sum_{i=1}^{n-1} u_i' \leq 1} (1-u_n)^{n-1} du_1' \cdots du_{n-1}' \right) du_n \qquad \text{transform } u_i' = \frac{u_i}{1-u_n} \\
&= \left(\int_{\sum_{i=1}^{n-1} u_i' \leq 1} du_1' \cdots du_{n-1}' \right) \left(\int_{u_n \leq 1} (1-u_n)^{n-1} du_n \right) \\
&= V_{n-1} \cdot \frac{1}{n}
\end{aligned}$$

We can apply this result recursively to get $V_n = \frac{1}{n!}$. Hence,

$$\mathbb{E}N = 1 + \sum_{n=1}^{\infty} \frac{1}{n!} = e$$

□

Example 3.3.3. Find $\max_{X \in [0,1]} \text{Var}(X)$. That is, what is the maximum possible variance of a RV X taking on values in the interval $[0,1]$?

Solution. Let $Y = 2X - 1$. Then

$$\max_{X \in [0,1]} \text{Var}(X) = \max_{Y \in [-1,1]} \text{Var}\left(\frac{Y+1}{2}\right) = \frac{1}{4} \max_{Y \in [-1,1]} \text{Var}\left(Y\right)$$

Note that $\text{Var}(Y) = \mathbb{E}Y^2 - (\mathbb{E}Y)^2 \leq \mathbb{E}Y^2 \leq 1$. To show the upper bound is attainable, we can construct $\mathbb{P}(Y = \pm 1) = \frac{1}{2}$, resulting in $\text{Var}(Y) = 1$. Hence, $\max_{Y \in [-1,1]} \text{Var}\left(Y\right) = 1$ and $\max_{X \in [0,1]} \text{Var}(X) = \frac{1}{4}$. □

3.4. Additional Problems

1. (a) Let $M_X(t)$ be the MGF of X, and define $K(t) = \log(M_X(t))$. Show that $K'(0) = \mathbb{E}X$ and $K''(0) = \operatorname{Var}(X)$.

   (b) Let

$$M_X(t) = \frac{(2\ell-1)!!}{(2\ell)!!} \frac{(t+2)(t+4)(t+6)\cdots(t+2\ell)}{(t+1)(t+3)(t+5)\cdots(t+2\ell-1)}$$

   where we define

$$r!! = \begin{cases} r(r-2)\cdots 5\cdot 3\cdot 1, & r \text{ odd} \\ r(r-2)\cdots 6\cdot 4\cdot 2, & r \text{ even} \end{cases}$$

   Show that M_X is a valid MGF (that is, $M_X(0) = 1$). Then, find an approximation for $\mathbb{E}X$ and $\operatorname{Var}(X)$ in terms of ℓ. The following information may be helpful:

$$\sum_{k=1}^{m} \frac{1}{k} \doteq \log(m) + \gamma + \frac{1}{2m} \qquad \text{and} \qquad \sum_{k=1}^{m} \frac{1}{k^2} \doteq \frac{\pi^2}{6} - \frac{1}{m}$$

   where $\gamma \doteq 0.5772$, known as the Euler-Mascheroni constant.

2. Let X be any non-negative RV. Prove that for any $r > 1$ we have

$$\int_0^\infty \frac{\mathbb{E}(X \wedge y^r)}{y^r} dy = \frac{r}{r-1} \mathbb{E} X^{1/r}$$

where $a \wedge b = \min(a, b)$.

3. Suppose $X_1, \cdots, X_{100} \sim F_X$, where F_X is continuous and strictly increasing. Let $\xi_1 = X_1$ and $\xi_k = X_k \wedge \xi_{k-1}$, where $a \wedge b = \min(a, b)$. Find the expected number of distinct elements in the set $\{\xi_1, \cdots, \xi_{100}\}$.
   Remark: There is a nice real-world interpretation to this problem. Suppose drivers $1, \cdots, 100$ are on a single-lane road, with driver 1 in the front, and let X_i by their desired speed to drive at. However, this desired speed may not be realized, since the driver in front might be driving slower, and hence our realized speed is in fact $\xi_k = X_k \wedge \xi_{k-1}$. Then, the number of distinct elements in $\{\xi_1, \cdots, \xi_{100}\}$ correspond to the number of pockets or groups of vehicles traveling together.

4. We denote $t^+ = \max(t, 0)$ and $t^- = \max(-t, 0)$. Let X, Y be random variables such that $\mathbb{E}X^+/\mathbb{E}X^- \geq z$ and $\mathbb{E}Y^+/\mathbb{E}Y^- \geq z$ for some $z > 0$. Does it follow that $\mathbb{E}(X+Y)^+/\mathbb{E}(X+Y)^- \geq z$?