7.1 Properties of Expectation

Recall:
- \( E(X) = \sum x \cdot p(x) \) for discrete \( X \),
- \( E(X) = \int x \cdot f(x) \, dx \) for continuous \( X \).

\[ E[g(X)] = \sum g(x) \cdot f(x) \, dx \]

\( E[g(x,y)] = \int \int g(x,y) \cdot f(x,y) \, dx \, dy \)

---

Two people agree to meet btw 12:00 and 1:00 pm.

They arrive at random times, with a uniform distribution in this interval.

**Expected waiting time?**

\( x, y \sim \text{Unif}(0,1) \) indep, \( E[|x-y|] = ? \)

\( f(x,y) = 1, \quad 0 \leq x, y \leq 1 \).

\[ E[|x-y|] = \int_0^1 \int_0^1 |x-y| \, dx \, dy \]

\[ = \int_0^1 \int_0^{y} (y-x) \, dx \, dy + \int_0^1 \int_y^1 (x-y) \, dx \, dy \]

\[ = \int_0^1 \frac{y^2}{2} \, dy + \int_0^1 \frac{(1-y)^2}{2} \, dy = \int_0^1 y^2 \, dy = \frac{1}{3} \]

\[ = \frac{20}{12} = \frac{5}{3} \text{ min} \]

---

\( \text{Ex.} \)

A dart is thrown at a round board of radius 1.

**Compute the expected dist. to the center.**

\( (X,Y) \sim \text{Unif}(0,1) \text{ circle} \)

\[ E[X^2+Y^2] = \int_0^1 \int_0^1 \sqrt{x^2+y^2} \, dx \, dy \rightarrow \text{polar coord:} \]

\[ = \frac{1}{2} \int_0^{2\pi} \int_0^1 r^3 \, dr \, d\theta = \frac{2\pi}{12} \int_0^1 r^3 \, dr = 2 \cdot \frac{r^4}{4} \bigg|_0^1 = \frac{2}{3} \]

---

\( \text{Alternative solution: by IIW, } \frac{\pi \cdot 4}{2} = 6.52, \text{ the dist. of a random pt is } \text{polar coord:} \)

\[ f(r, \theta) = \frac{1}{2\pi}, \quad 0 \leq r \leq 1, \quad 0 \leq \theta < 2\pi \]

\[ \phi_r(r) = \frac{1}{2\pi}, \quad 0 \leq r \leq 1 \]

\[ E[\phi_r] = \int_0^1 r^2 \, dr = \frac{r^3}{3} \bigg|_0^1 = \frac{2}{3} \]
\[ \forall X, Y: \quad E(X + Y) = E(X) + E(Y) \]

- Not even independent!
- We formulated and used this before, but never proved.

\[ E(X + Y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x + y) f(x, y) \, dx \, dy \]
\[ = \int_{-\infty}^{\infty} (\int_{-\infty}^{\infty} f(x, y) \, dy) \, dx + \int_{-\infty}^{\infty} f_x(x) \, dx \]
\[ = \int_{-\infty}^{\infty} x f_x(x) \, dx + \int_{-\infty}^{\infty} y f_y(y) \, dy = E(X) + E(Y). \]

**Cor**
\[ E(\sum_i X_i) = \sum_i E(X_i) \]

(by induction)

**Prop**
\[ X \geq 0 \Rightarrow E(X) \geq 0 \]

(for cont.)

\[ E(X) = \int_{\mathbb{R}} x f(x) \, dx \geq 0 \]

**Cor**
\[ X > Y \Rightarrow E(X) = E(Y) \]

\[ X - Y > 0 \Rightarrow E(X - Y) > 0 \]
\[ E(X - Y) = E(X) - E(Y). \]

**Cor**
\[ a < X < b \Rightarrow a < E(X) < b \]

(follows from previous cor.)
7.2 - 7.3. Prop Expectation of sums of r.v.'s

\[ E(x) = E(x_1) + \cdots + E(x_n) \]

\[ E(x) \quad \text{even if not indep} \]

Application to counting questions

**Example (Deaths in the town [D. Bernoulli 1720-1782])** - [Ghahramani Ex. 10.3]

- \( n \) married couples live in a town.
- \( n \) death occur at random in the town.
- Expected \( k \) of intact couples?

\[ X = X_1 + \cdots + X_n \quad \text{where} \quad X_i = \begin{cases} 1, & \text{\( i \)'th couple is intact} \\ 0, & \text{otherwise} \end{cases} \]

\[ X_i \sim \text{Bernoulli}(p). \]

\[ p = P(X_i = 1) = P(\text{\( i \)'th couple is intact}) \]

\[ = P(\text{all but the \( i \)'th death occur among the other \( n-1 \) couples}) \]

\[ = \frac{\binom{2n-2}{m}}{\binom{2n}{m}} = \frac{(2n-2)!}{m! (2n-m-1)!} \cdot \frac{1}{(2n-2)} \cdot \frac{1}{(2n-1)} \]

\[ = \frac{(2n-m)(2n-m-1)}{2n 3(2n-1)} \]

\[ EX = \sum_{i=1}^{n} E(X_i) = n \cdot E(X_1) = n \cdot p = \frac{(2n-m)(2n-m-1)}{2(2n-1)}. \]

Illustration: \( n=1000 \)

<table>
<thead>
<tr>
<th>( m )</th>
<th>( E(x) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>90.2</td>
</tr>
<tr>
<td>600</td>
<td>490</td>
</tr>
<tr>
<td>1200</td>
<td>150</td>
</tr>
<tr>
<td>1500</td>
<td>62</td>
</tr>
<tr>
<td>1800</td>
<td>10</td>
</tr>
</tbody>
</table>
Ex. (Coupon Collecting Problem)

There are $n$ coupons of different types of coupons. Each time one obtains a coupon, it is equally likely to be of any type.

Compute the expected number of the different coupons among $N$ collected.

Applications: Clinical trials—waits for side effects of drug.

Let $Y = N - X$, where $X$ is the number of uncollected coupons.

**$E[Y] = ?$**

$Y = Y_1 + Y_2 + \ldots + Y_n$, where $Y_i = \begin{cases} \frac{1}{n}, & \text{coupon of } i^{th} \text{ type is not collected} \\ 0, & \text{otherwise} \end{cases}$

$$E[Y] = \sum_{i=1}^{n} E[Y_i] = n \cdot \frac{1}{n} \cdot 0 = n \cdot \frac{1}{n} = 1.$$ 

$p = P\{\text{coupon of } i^{th} \text{ type is not collected}\} = \left(1 - \frac{1}{n}\right)^N$. 

So $E[Y] = \sum_{i=1}^{N} p = n \cdot \left(1 - \frac{1}{n}\right)^N$.

$$E[X] = n - n \cdot \left(1 - \frac{1}{n}\right)^N$$

Asymptotic Analysis: $n \to \infty$, $N = \ln n$

$$E[Y] \approx n \cdot e^{-N/n} = n \cdot \frac{ne^{-t}}{n!}.$$ 

Then $E[Y] < 1$ for $t \sim \ln n \Rightarrow$ for $N \sim \ln n$.

Should expect a complete collection in time $N \sim \ln n$.

Let's verify this:

---

-90-