Section 12

Wednesday, November 16, 2022 10:57 PM

Exchangeability

Dealing w/ multiple r.v.'s. – sequences of them ...

(Def) A collection of R.V.'s is exchangeable if $(X_1, X_2 \ldots X_n) \overset{\mathcal{D}}{=} (X_{i_1}, X_{i_2} \ldots X_{i_n})$, where $(i_1, i_2 \ldots i_n)$ is a permutation of $(1, 2 \ldots n)$.

(Def) $f: \mathbb{R}^n \to \mathbb{R}$ is a symmetric function if $f(x_1 \ldots x_n) = f(x_{i_1} \ldots x_{i_n})$, where $\{i_j\}_{j=1}^n$ is as above.

(Lm) A collection $X_1 \ldots X_n$ of r.v.'s is exchangeable if and only if their joint pmf/pdf is a symmetric function.

(Th) If $X_1 \ldots X_n$ are exchangeable, then for all $1 \le k \le n$,
$(X_1 \ldots X_k) \overset{\mathcal{D}}{=} (X_{i_1} \ldots X_{i_k})$, where $\{i_j\}_{j=1}^k$ is a subset of $\{1 \ldots n\}$ of length k.

* iid r.v.'s are exchangeable, but not all exchangeable r.v.'s are iid.

(Q) You draw randomly (100 students) & uniformly, with replacement, students from a sample of 100.
Let Z_i = index of student you draw on i^{th} trial.
Compute:

$S_1 \; S_2 \; S_3 \; \ldots \; S_{100}$

(1) $\mathbb{P}[Z_1 = 1]$
(2) $\mathbb{P}[Z_2 = 1]$
(3) $\mathbb{P}[Z_1 = 1, Z_2 = 2]$
(4) $\mathbb{P}[Z_8 = 1, Z_{99} = 2]$
(5) $\mathbb{P}[Z_3 = 1, Z_4 = 3, Z_9 = 2]$
(6) $\mathbb{P}[Z_8 = 9, Z_{11} = 9]$

(Q) Same as above, but now you draw without replacement. What changes? → think:
(1) identically distributed?
(2) independent?
(3) exchangeable?

Now, prove that $Z_1, \ldots Z_{100}$ are exchangeable.

(Q) Say, $X_1, X_2 \,\&\, X_3 \sim \mathcal{N}(0,1)$ iid.
Compute:
(1) $\mathbb{P}[X_1 > X_2]$
(2) $\mathbb{P}[X_1 > X_2 > X_3]$
(3) $\mathbb{P}[X_3 > X_1]$

Linearity of expectation

(Lm) $\mathbb{E}[\sum X_i] = \sum \mathbb{E}[X_i]$

New notion - covariance

(Df) $\mathrm{Cov}(X,Y) := \mathbb{E}[(X-\mu_X)(Y-\mu_Y)]$
$= \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]$

(Lm) Independence $\Rightarrow$ uncorrelated
Uncorrelated $\not\Rightarrow$ Independence.

(Th) We have: Properties following defn

(i) $\mathrm{Cov}(X,Y) = \mathrm{Cov}(Y,X)$

(ii) $\mathrm{Cov}(X,X) = \mathrm{Var}(X)$

(iii) $\mathrm{Cov}(\sum a_i X_i, Y) = \sum a_i \mathrm{Cov}(X_i, Y)$

$\mathrm{Cov}(X, \sum a_i Y_i) = \sum a_i \mathrm{Cov}(X, Y_i)$

"Linearity" of Variance?

$$\mathrm{Var}\left(\sum_{i=1}^{n} a_i X_i\right) = \mathrm{Cov}\left(\sum_{i=1}^{n} a_i X_i, \sum_{j=1}^{n} a_j X_j\right)$$
$$= \sum a_i \mathrm{Cov}(X_i, \sum a_j X_j)$$
$$= \sum\sum a_i a_j \mathrm{Cov}(X_i, X_j)$$
$$= a_i^2 \sum \mathrm{Var}(X_i) + \sum_{i \neq j}\sum a_i a_j \mathrm{Cov}(X_i, X_j)$$

Correlation

$$\rho_{X,Y} := \frac{\mathrm{Cov}(X,Y)}{\sigma_X \sigma_Y}$$

Cauchy - Schwarz

If X,Y are r.v's w/ finite mean & variance,
$|\mathbb{E}(XY)| \le \sqrt{\mathbb{E}(X^2)}\sqrt{\mathbb{E}(Y^2)}$
and
$|\mathrm{Cov}(X,Y)| \le \sqrt{\mathrm{Var}(X)}\sqrt{\mathrm{Var}(Y)}$

Example 2.2.7. Coupon collector's problem. Let $X_1, X_2, \ldots$ be i.i.d. uniform on $\{1, 2, \ldots, n\}$. To motivate the name, think of collecting baseball cards (or coupons). Suppose that the ith item we collect is chosen at random from the set of possibilities and is independent of the previous choices. Let $\tau_k^n = \inf\{m : |\{X_1, \ldots, X_m\}| = k\}$ be the first

60 *CHAPTER 2. LAWS OF LARGE NUMBERS*

time we have k different items. In this problem, we are interested in the asymptotic behavior of $T_n = \tau_n^n$, the time to collect a complete set. It is easy to see that $\tau_1^n = 1$. To make later formulas work out nicely, we will set $\tau_0^n = 0$. For $1 \le k \le n$, $X_{n,k} \equiv \tau_k^n - \tau_{k-1}^n$ represents the time to get a choice different from our first $k-1$, so $X_{n,k}$ has a geometric distribution with parameter $1 - (k-1)/n$ and is independent of the earlier waiting times $X_{n,j}$, $1 \le j < k$. Example 1.6.14 tells us that if X has a geometric distribution with parameter p then $EX = 1/p$ and $\mathrm{var}(X) \le 1/p^2$. Using the linearity of expected value, bounds on $\sum_{m=1}^n 1/m$ in (2.2.1), and Theorem 2.2.1 we see that

$$ET_n = \sum_{k=1}^n \left(1 - \frac{k-1}{n}\right)^{-1} = n \sum_{m=1}^n m^{-1} \sim n \log n$$
$$\mathrm{var}(T_n) \le \sum_{k=1}^n \left(1 - \frac{k-1}{n}\right)^{-2} = n^2 \sum_{m=1}^n m^{-2} \le n^2 \sum_{m=1}^\infty m^{-2}$$

Taking $b_n = n \log n$ and using Theorem 2.2.6, it follows that

$$\frac{T_n - n\sum_{m=1}^n m^{-1}}{n \log n} \to 0 \quad \text{in probability}$$

and hence $T_n/(n \log n) \to 1$ in probability.

For a concrete example, take $n = 365$, i.e., we are interested in the number of people we need to meet until we have seen someone with every birthday. In this case the limit theorem says it will take about $365 \log 365 = 2153.46$ tries to get a complete set. Note that the number of trials is 5.89 times the number of birthdays.

(Q) Say, $X_1 \ldots X_n \sim$ iid $N(\mu, \sigma^2)$.
Knowing that the sum of normal r.v's is a normal r.v.,
consider $S_n = \sum X_i$, "sample total". Calculate:

(1) $\mathbb{E}[S_n]$

(2) $\mathrm{Var}[S_n]$

(3) $\mathbb{P}[\frac{S_n}{n} > \mu]$

I strongly recommend you read and try out the examples in pages 240-243 of Gabe's book

Exploiting linearity of expectation

I strongly recommend you read and try out the examples in pages 240-243 of Gabe's book

365 log 365 = 2153.46 tries to get a complete set. Note that the number of trials is 5.89 times the number of birthdays.

Exploiting linearity of expectation

Coupon collector

Say, you are collecting n coupons. Let T_n = time taken to collect all n.

Compute $\mathbb{E}[T_n]$ and $Var(T_n)$ if:

(1) $x = 1$

(2) $x = 5$

Record values

Say, you walk from your house to Ames for section every thursday, for 14 Thursdays. Let X_i be the time taken each Thursday to reach class. $X_i \sim \exp(\frac{1}{10})$ iid.

You are interested in breaking your 'record', i.e. arriving to class faster than you did on any previous day.

(1) What does $\{X_2 > X_1\}$ denote? What is $\mathbb{P}[X_2 > X_1]$?
What is $\mathbb{P}[X_3 > X_2 > X_1]$?

(2) Suppose you're interested in breaking a record on the i^{th} trial.
Let A_i be the event that $X_i >$ all previous X_j's.
-> If $i = 3$, what is $P[A_i]$?
-> For general i, what is $P[A_i]$?

(3) Suppose you're interested in counting the number of times you break the record. Let N represent this.

(a) How can you express N in terms of A_i?

(b) What is $\mathbb{E}[N]$?

(4) Say, you're interested in how long you must wait to exceed X_1. Let N be this r.v.

What are we even doing right now?!

So far into the course, we have used everyday probability to motivate the usage of random variables, and then dived into the study of random variables in the discrete and continuous case (defined as having either countable and uncountable sample spaces, respectively). In the latter part, we started to think about the distribution of two or more random variables - joint distributions.

Now, what we are going to do is something even more powerful. We are going to think about scenarios that are best modelled by multiple random variables - a **finite sequence of random variables,** $X_1 ... X_n$, to be precise. As it will turn out, it is often that such sequences will possess useful properties that will enable us to understand their 'amalgamated' random variable, for ex:

- $\sum X_i$, sum or total
- $\overline{X} := \frac{\sum X_i}{n}$, sample mean

Particular properties that we will try to look out for are:

- Independence
- Identically Distributed
- Exchangeability