

Prelims Probability extra questions

Leo Tyrpak

September 2024

1 Sums of random variables

Suppose X_1, \dots, X_n are independent and identically distributed (i.i.d.) with $\mathbf{E}[X_1] = 0$. Let $S_n = X_1 + \dots + X_n$.

Suppose $\mathbf{E}[|X_1|^k] < \infty$ for $k \in \mathbb{N}$. What is $\mathbf{E}[S_n^k]$? What is $M_{S_n}(t)$ in terms of $M_{X_1}(t)$?

2 Combinatorics

1. For sheet 8 question 7, what happens if $p = a$ or $p = b$? What does the limit approach, why?
2. Let U_1, \dots, U_n i.i.d Uniform $[0,1]$ random variables. What is $\mathbf{P}\{U_1 + \dots + U_n \leq 1\}$?
3. During a game of poker you are dealt a hand of 5 cards uniformly at random. With the convention that aces can be either high or low for a straight. What is the probability of having 1 pair, 2pairs, 3 of a kind, straight, flush, full house, 4 of a kind, straight flush. See Poker hands for a list of poker hands and what they mean.
4. 8 pawns are placed uniformly at random on a chess board, no more than one per square. What is the probability that no 2 are in the same row or column?
5. Players A, B, C throw a uniform dice at random in the order ABCAB-CABC... Each dice throw is independent. What is the probability that A throws a 6 first? What is the probability that A throws a 6 first and B throws a 6 the second and C is the third to throw a 6?
6. Each member of a group of n players roll a dice. For any pair of players who roll the same number, the group scores 1 point. Find the mean and variance of the total score of the group. Find the mean and variance of the total score if any pair of players who throw the same number scores that number.

3 Probabilistic method

This is also sometimes called probabilistic combinatorics.

Let $G=(V,E)$ be a finite simple graph (so no loops or multiple edges between 2 vertices). Write d_v for the degree of a vertex v . Prove the following:

1. An independent set is a set of vertices no 2 of which share an edge. Denote by $\alpha(G)$ the largest independent set. Prove that:

$$\alpha(G) \geq \sum_v \frac{1}{d_v + 1} \quad (1)$$

2. For any set $W \subseteq V$, edge $e \in E$:

$$1_W(e) = 1 \text{ if } e \text{ connects } W \text{ and } W^c \quad (2)$$

If $N_W = \sum_e 1_W(e)$ then prove that there exists W such that $N_W \geq 1/2|E|$

We construct a random graph G' by including each edge independently with probability $p \in (0, 1)$.

4 Analysis

1. Prove Bonferroni's inequality

$$\mathbf{P} \left\{ \bigcup_{r=1}^n A_r \right\} \geq \sum_{r=1}^n \mathbf{P} \{A_r\} - \sum_{r < k} \mathbf{P} \{A_r \cap A_k\} \quad (3)$$

2. Prove Kounia's inequality

$$\mathbf{P} \left\{ \bigcup_{r=1}^n A_r \right\} \leq \min_k \left\{ \sum_{r=1}^n \mathbf{P} \{A_r\} - \sum_{r \neq k} \mathbf{P} \{A_r \cap A_k\} \right\} \quad (4)$$

4.1 Moment generating functions

Sheet 3 Question 5 asks you to prove that,

$$\mathbf{E} [\exp(tX)] = \exp(\lambda(e^t - 1))$$

for X a Poisson random variable of parameter λ .

The expression on the left is called a moment generating function(mgf) and is a powerful analytical tool useful in probability to get concentration inequalities and uniqueness of distributions.

We write $M_X(t) = \mathbf{E} [\exp(tX)]$ for the mgf at t .

Calculate the mgf of the following:

1. Geo(p)

2. Ber(p)
3. Uniform on $\{1, \dots, n\}$
4. Bin(n,p)
5. Exponential rate λ
6. Normal random variable mean μ , variance σ^2

Let a, b be some constants, then what is the mgf of $Z_1 = aX + b$ (i.e. $M_{Z_1}(t)$)?

Now suppose X, Y are independent, then what is the mgf of $Z_2 = X + Y$ (i.e. $M_{Z_2}(t)$)?

You might have noticed that the mgf doesn't always exist, as sometimes the expectation of $\exp(tX)$ is infinite. Can you construct an example of a random variable where $\mathbf{E}[\exp(tX)]$ is infinite for all $t \neq 0$?

By differentiating (and not worrying about the important analysis that needs to be done in the background but that you will only do later), can you relate the moments of a distribution with the mgf?

Note: a moment is an expression of the form $\mathbf{E}[X^k]$. First moment is mean $\mathbf{E}[X]$ and second moment is $\mathbf{E}[X^2]$.

Use this to find expressions for the mean and variance of a random variable X using the mgf. Then check this with the distributions considered above.

Prove that for all t :

$$\mathbf{P}\{X > x\} \leq M_X(t) \exp(-tx)$$

Consider X is Ber(p), Geo(p), Bin(n,p), Poi(λ), Exp(λ) and for each of those optimise the right hand side in t to get optimal tail bounds.

4.2 Probability generating functions

Sheet 5 Question 4 asks you to calculate the probability generating function of a Geometric distribution.

Calculate the following pgf:

1. Ber(p)
2. Bin(n,p)
3. Poisson(λ)
4. Uniform $\{1, \dots, n\}$
5. Negative Binomial (r,p) with pdf $p_k = \binom{k+r-1}{k} (1-p)^k p^r$

Let a, b be some constants, then what is the pgf of $Z_1 = aX + b$ (i.e. $M_{Z_1}(t)$)?

Now suppose X, Y are independent, then what is the pgf of $Z_2 = X + Y$ (i.e. $M_{Z_2}(t)$)?

Can we recover the distribution of a random variable from the pgf?

5 Miscellaneous

5.1 Sheet 3 Question 6

Sheet 3 Question 6 asks us to find out how long we have to wait on average to see 2 heads in a row, then Sheet 5 Question 5 asks us to calculate the exact distribution.

1. How long do you have to wait on average to get k heads in a row where $k \in \mathbb{N}$?
2. What is $r_n = \mathbf{P}\{X > n\}$ where X is the first time you see k heads in a row?
3. Let Z be the first time you see HTH . What is $\mathbf{E}[Z]$ and $\mathbf{P}\{Z > n\}$?
4. Let Z be the first time you see $HTTH$. What is $\mathbf{E}[Z]$?
5. If the challenge doesn't prove enough, try to extend your methods to any sequence.