Part IA - Probability
Theorems with Proof

Lectured by R. Weber

Lent 2015

Basic concepts
Classical probability, equally likely outcomes. Combinatorial analysis, permutations and combinations. Stirling's formula (asymptotics for $\log n!$ proved). [3]

Axiomatic approach
Axioms (countable case). Probability spaces. Inclusion-exclusion formula. Continuity and subadditivity of probability measures. Independence. Binomial, Poisson and geometric distributions. Relation between Poisson and binomial distributions. Conditional probability, Bayes's formula. Examples, including Simpson's paradox. [5]

Discrete random variables
Expectation. Functions of a random variable, indicator function, variance, standard deviation. Covariance, independence of random variables. Generating functions: sums of independent random variables, random sum formula, moments.

Conditional expectation. Random walks: gambler's ruin, recurrence relations. Difference equations and their solution. Mean time to absorption. Branching processes: generating functions and extinction probability. Combinatorial applications of generating functions. [7]

Continuous random variables
Distributions and density functions. Expectations; expectation of a function of a random variable. Uniform, normal and exponential random variables. Memoryless property of exponential distribution. Joint distributions: transformation of random variables (including Jacobians), examples. Simulation: generating continuous random variables, independent normal random variables. Geometrical probability: Bertrand's paradox, Buffon's needle. Correlation coefficient, bivariate normal random variables. [6]

Inequalities and limits
Markov's inequality, Chebyshev's inequality. Weak law of large numbers. Convexity: Jensen's inequality for general random variables, AM/GM inequality.

Moment generating functions and statement (no proof) of continuity theorem. Statement of central limit theorem and sketch of proof. Examples, including sampling. [3]

Contents

1 Introduction

2 Classical probability

2.1 Classical probability

2.2 Sample space and events

3 Combinatorial analysis

3.1 Counting

Theorem (Fundamental rule of counting)**.** Suppose we have to make r multiple choices in sequence. There are m_1 possibilities for the first choice, m_2 possibilities for the second etc. Then the total number of choices is $m_1 \times m_2 \times \cdots m_r$.

3.2 Sampling with or without replacement

3.3 Sampling with or without regard to ordering

3.4 Four cases of enumerative combinatorics

4 Stirling's formula

4.1 Multinomial coefficient

4.2 Stirling's formula

Proposition. $\log n! \sim n \log n$

Proof. Note that

$$e^n = 1 + n + \frac{n^2}{2!} + \cdots + \frac{n^n}{n!} + \cdots.$$

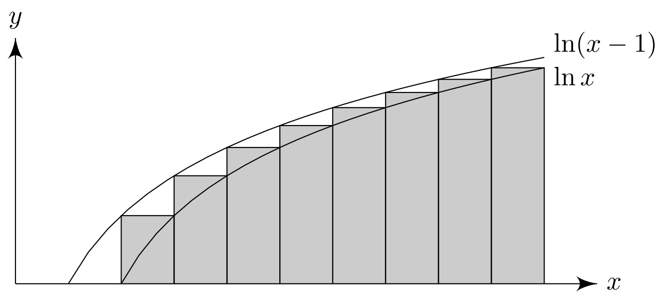
So $1 \leq n^n/n! \leq e^n$. Note that

$$\log n! = \sum_{k=1}^{n} \log k.$$

Now we claim that

$$\int_1^n \log x \, \mathrm{d}x \leq \sum_1^n \log k \leq \int_1^{n+1} \log x \, \mathrm{d}x.$$

This is true by considering the diagram:



Perform the integral to obtain

$$n \log n - n + 1 \leq \log n! \leq (n+1)\log(n+1) - n;$$

Divide both sides by $n \log n$ and let $n \to \infty$. Both sides tend to 1. So

$$\frac{\log n!}{n \log n} \to 1.$$

□

Theorem (Stirling's formula)**.** As $n \to \infty$,

$$\log\left(\frac{n!e^n}{n^{n+\frac{1}{2}}}\right) = \log\sqrt{2\pi} + O\left(\frac{1}{n}\right)$$

Corollary.

$$n! \sim \sqrt{2\pi} n^{n+\frac{1}{2}} e^{-n}$$

Proof. (non-examinable) Define

$$d_n = \log\left(\frac{n!e^n}{n^{n+1/2}}\right) = \log n! - (n+1/2)\log n + n$$

Then

$$d_n - d_{n+1} = (n+1/2)\log\left(\frac{n+1}{n}\right)$$

Write $t = 1/(2n+1)$. Then

$$d_n - d_{n+1} = \frac{1}{2t}\log\left(\frac{1+t}{1-t}\right).$$

We can simplifying by noting that

$$\begin{aligned}
\log(1+t) - t &= -\frac{1}{2}t^2 + \frac{1}{3}t^3 - \frac{1}{4}t^4 + \cdots \\
\log(1-t) + t &= -\frac{1}{2}t^2 - \frac{1}{3}t^3 - \frac{1}{4}t^4 - \cdots
\end{aligned}$$

Then if we subtract the equations and divide by $2t$, we obtain

$$\begin{aligned}
d_n - d_{n+1} &= \frac{1}{3}t^2 + \frac{1}{5}t^4 + \frac{1}{7}t^6 \\
&\leq \frac{1}{3}t^2 + \frac{1}{3}t^4 + \frac{1}{3}t^6 = \cdots \\
&= \frac{1}{3}\frac{t^2}{1-t^2} \\
&= \frac{1}{3}\frac{1}{(2n+1)^2 - 1} \\
&= \frac{1}{12}\left(\frac{1}{n} - \frac{1}{n+1}\right)
\end{aligned}$$

By summing these bounds, we know that

$$d_1 - d_n \leq \frac{1}{12}\left(1 - \frac{1}{n}\right)$$

Then we know that d_n is bounded below by d_1+ something, and is decreasing since $d_n - d_{n+1}$ is positive. So it converges to a limit A.

Suppose $m > n$. Then $d_n - d_m < \left(\frac{1}{n} - \frac{1}{m}\right)\frac{1}{12} - \frac{2}{15}\frac{1}{(2n+1)^4}$ by adding back the term we removed when changing $t^4/5$ to $t^4/3$. So taking the limit as $m \to \infty$, $A < d_n < A + 1/(12n)$, with strict inequalities due to the second term.

To find A, we have a small detour to prove a formula:

Take $I_n = \int_0^{\pi/2} \sin^n\theta \,\mathrm{d}\theta$. This is decreasing as n increases as $\sin^n\theta$ gets smaller. We also know that

$$\begin{aligned}
I_n &= \int_0^{\pi/2} \sin^n\theta \,\mathrm{d}\theta \\
&= -\cos\theta\sin^{n-1}\theta|_0^{\pi/2} + \int_0^{\pi/2}(n-1)\cos^2\theta\sin^{n-2}\theta \,\mathrm{d}\theta \\
&= (n-1)(I_n - 2 - I_n)
\end{aligned}$$

So

$$I_n = \frac{n-1}{n} I_{n-2}.$$

We can directly evaluate the integral to obtain $I_0 = \pi/2$, $I_1 = 1$. Then

$$\begin{aligned} I_{2n} &= \frac{1}{2} \cdot \frac{3}{4} \cdots \frac{2n-1}{2n} \pi/2 = \frac{(2n)!}{(2^n n!)^2} \frac{\pi}{2} \\ I_{2n+1} &= \frac{2}{3} \cdot \frac{4}{5} \cdots \frac{2n}{2n+1} = \frac{(2^n n!)^2}{(2n+1)!} \end{aligned}$$

So using the fact that I_n is decreasing, we know that

$$1 \leq \frac{I_{2n}}{I_{2n+1}} \leq \frac{I_{2n-1}}{I_{2n+1}} = 1 + \frac{1}{2n} \to 1.$$

Using the approximation $n! \sim n^{n+1/2} e^{-n+A}$, where A is the limit we want to find, we can approximate

$$\frac{I_{2n}}{I_{2n+1}} = \pi(2n+1)\left[\frac{((2n)!)^2}{2^{4n+1}(n!)^4}\right] \sim \pi(2n+1)\frac{1}{ne^{2A}} \to \frac{2\pi}{e^{2A}}.$$

Since the last expression is equal to 1, we know that $A = \log\sqrt{2\pi}$ $\square$

Proposition ((non-examinable))**.** We use the $1/12n$ term from the proof above to get a better approximation:

$$\sqrt{2\pi} n^{n+1/2} e^{-n} + \frac{1}{12n+1} \leq n! \leq \sqrt{2\pi} n^{n+1/2} e^{-n} + \frac{1}{12n}.$$

5 Axiomatic approach

Theorem.

(i) $P(\emptyset) = 0$

(ii) $P(A^C) + 1 - P(A)$

(iii) $A \subseteq B \Rightarrow P(A) \leq P(B)$

(iv) $P(A \subseteq B) = P(A) + P(B) - P(A \cap B)$.

(v) Let $A_1 \subseteq A_2 \subseteq A_3 \subseteq \cdots$. Then

$$P\left(\bigcup_1^\infty A_i\right) = \lim_{n\to\infty} P(A_n).$$

This states that P is a continuous set function.

Proof.

(i) Ω and $\emptyset$ are disjoint. So $P(\Omega) + P(\emptyset) = P(\Omega \cup \emptyset) = P(\Omega)$. So $P(\emptyset) = 0$.

(ii) $P(\Omega) = 1 = P(A) + P(A^C)$ since A and A^C are disjoint.

(iii) Write $B = A \cup (B \cap A^C)$. Then $P(B) = P(A) + P(B \cap A^C) \geq P(A)$.

(iv) $P(A \cup B) = P(A) + P(B \cap A^C)$. We also know that $P(B) = P(A \cap B) + P(B \cap A^C)$. Then the result follows.

(v) c.f. Lecture 7

□

5.1 Boole's inequality

Theorem (Boole's inequality)**.** For any $A_1, A_2, \cdots$,

$$P\left(\bigcup_{i=1}^\infty A_i\right) \leq \sum_{i=1}^\infty P(A_i).$$

Proof. The axiom states a similar formula that only holds for disjoint sets. So we need a clever trick to make them disjoint. We define

$$\begin{aligned} B_1 &= A_1 \\ B_2 &= A_2 \setminus B_1 \\ B_i &= A_i \setminus \bigcup_{k=1}^{i-1} A_k. \end{aligned}$$

So we know that

$$\bigcup B_i = \bigcup A_i.$$

But the B_i are disjoint. So our Axiom (iii) gives

$$P(\bigcup_i A_i) = P(\bigcup_i B_i) = \sum_i P(B_i) \leq \sum_i P(A_i).$$

Where the last inequality follows from (iii) of the theorem above. □

5.2 Inclusion-exclusion formula

Theorem (Inclusion-exclusion formula)**.**

$$P\left(\bigcup_i^n A_i\right) = \sum_1^n P(A_i) - \sum_{i_1<i_2} P(A_{i_1} \cap A_{j_2}) + \sum_{i_1<i_2<i_3} P(A_{i_1} \cap A_{i_2} \cap A_{i_3}) - \cdots$$
$$+ (-1)^{n-1} P(A_1 \cap \cdots \cap A_n).$$

Proof. Perform induction on n. $n = 2$ is proven above.

Then

$$P(A_1 \cup A_2 \cup \cdots A_n) = P(A_1) = P(A_2 \cup \cdots \cup A_n) - P\left(\bigcup_{i=2}^n (A_1 \cap A_i)\right).$$

and we can apply the induction hypothesis for $n - 1$. □