Probability and Other Branches of Mathematics

Jordan Stoyanov (Newcastle/Ljubljana)

e-mail: stoyanovj@gmail.com

Berliner Mathematische Kolloquium

TU, Berlin, GERMANY

11 November 2015, 18:15

Goal Today:

Discussion on: real and integer numbers, combinatorial identities, functions, integrals, infinite series, systems of functional or algebraic equations, ...

We need: Some combinatorics, algebra, analysis, number theory, ...

Little Probability: random events, random variables, probabilities, distributions, densities, moments, ... , also the LLN.

Standard Notations: $\binom{n}{k}$, ... , $\mathbb{N}$, $\mathbb{R}$, ... , $\mathcal{N}$, ..., $\mathbf{P}$, $\mathbf{E}$, m_k.

Intriguing/Strange questions ... followed by **Beautiful answers.**

Maybe Surprising: The use of ideas/techniques from Probability.

Modern Mathematics: Long story. Areas. For good and for bad.

Directly and/or indirectly mutually interrelated: ideas, notions, results, techniques, consequences, etc., by keeping specifics!

History: Theory of Probability ... Kolmogorov (1933), based on Borel, Lebesgue, Daniel, Caratheodori, ... Sceptics like Hardy also contributed, without knowing ... [Recent **Hardy's condition ....**]

Probability, Analysis, Discrete Math: Natural ideas to exploit: Probabilities, Expectations are real numbers, density integrates to 1, etc.

Probabilistic Method: Paul Erdös and Alfred Rényi: In a finite set of objects $\mathcal{A}_n$, existence of an object with a given property α. If $\mathbf{P}[\text{no } \alpha] > 0$.

Striking Cases: Graph theory, Number theory, PDEs (Krylov - Evans).

Warning: Any serious area of research has to be taken seriously!

Discrete Uniform Distribution: Random variable (r.v.) $X \sim U_n$.
The value of X is a number in the set $\{1, 2, \ldots, n\}$, each has probab $\frac{1}{n}$.

Die (fair, standard, symmetric): each of $1, 2, 3, 4, 5, 6$, probab $\frac{1}{6}$.

Continuous Uniform on $(0, 1)$: $X \sim U(0, 1)$, values in $(0, 1)$, density $f(x) = 1$ for $x \in (0, 1)$; 0 otherwise. $U(0, 1) = \beta(1, 1)$, what is $\beta(a, b)$?

Beta distribution: $X \sim \beta(a, b)$, $a > 0, b > 0$, X continuous, density

$$f(x) = \frac{x^{a-1}(1-x)^{b-1}}{B(a, b)}, \ x \in (0, 1).$$

Recall that $B(a, b)$ and $\Gamma(a)$ are the Euler beta- and gamma-function:
$B(a, b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$, $\Gamma(a) = \int_0^\infty x^{a-1}e^{-x}\mathrm{d}x$; $\Gamma(n+1) = n!$.

Combinatorial Identity: For any $k, n \in \mathbb{N}$, $1 \le k \le n$, one holds:

$$\binom{n}{k} \sum_{j=0}^{k} (-1)^{k-j} \frac{n+1}{n+1-j} \binom{k}{j} = 1. \qquad (*)$$

It looks $\sim$ easy, as it is!

Proof: Use that: Density $f(x), x \in I \iff f \ge 0$ on I and $\int_I f(x)\mathrm{d}x = 1$.

Take a r.v. $X \sim \beta(n-k+1, k+1)$, so its density is

$$f(x) = \frac{x^{n-k}(1-x)^k}{B(n-k+1, k+1)}, \ x \in (0,1).$$

Now, use Newton's binomial formula for $(1-x)^k$, the fact that $\Gamma(n+1) = n!$, and see that

$$
\begin{aligned}
1 &= \int_0^1 f(x)\mathrm{d}x = \frac{1}{B(n-k+1,k+1)} \int_0^1 x^{n-k}(1-x)^k \mathrm{d}x \\
&= \frac{\Gamma(n+2)}{\Gamma(n-k+1)\Gamma(k+1)} \int_0^1 x^{n-k} \sum_{j=0}^{k} \binom{k}{j} (-1)^{k-j} x^{k-j} \mathrm{d}x \\
&= \frac{\Gamma(n+2)}{\Gamma(n-k+1)\Gamma(k+1)} \sum_{j=0}^{k} \binom{k}{j} (-1)^{k-j} \int_0^1 x^{n-j} \mathrm{d}x \\
&= \frac{(n+1)!}{(n-k)!\,k!} \sum_{j=0}^{k} (-1)^{k-j} \binom{k}{j} \frac{1}{n-j+1} = (*).
\end{aligned}
$$

Exercise: Find another proof of (*).

Two Identities more:

First: Here $n,\ k,\ j \in \mathbb{N}$, with $j \le k \le n$ and $s \in \mathbb{N}_0$, Show that

$$\sum_{m=j}^{n-k+j} \binom{m-1}{j-1}\binom{n-m}{k-j} = \binom{n}{k}; \qquad \sum_{m=j}^{n-k+j} \binom{m+s-1}{j+s-1}\binom{n-m}{k-j} = \binom{n+s}{k+s}.$$

Story: IMO 1980, Washington. BG team, 8 full solutions.

Second: For any $n \in \mathbb{N}$,

$$\sum_{i=1}^{n} \binom{2i}{i}\binom{2n-2i}{n-i} = 4^n.$$

Hint: Uses $\mathbf{E}[(X_1^2 + X_2^2)^2]$, $X_1,\ X_2 \sim \mathcal{N}(0,1)$, indep. $(X^2 \sim \chi_1^2)$

Hadamard Inequality: If $f(x),\ x \in (a,b)$ is a convex function, then

$$f\left(\tfrac{a+b}{2}\right) \;\le\; \tfrac{1}{a+b} \int_a^b f(t)\,\mathrm{d}t \;\le\; \tfrac{1}{2}\left(f(a)+f(b)\right). \qquad (*)$$

Hint: Convexity $\Rightarrow\ f(x) \ge f(u) + (x-u)f'(u),$ any $u, x \in (a,b)$.

Take r.v. $X \sim F$ on (a,b), by Jensen Ineq. and $x = \mathbf{E}X,\ u = X$, find

$$0 \le \mathbf{E}[f(X)] - f(\mathbf{E}X) \le \mathbf{E}[(X-\mathbf{E}X)\,f'(X)].$$

Equality iff f is linear, or X takes one fixed value w.p. 1.

Then, let $X \sim$ Uniform on (a,b), density $\frac{1}{b-a}$ on (a,b), $\mathbf{E}X = \frac{1}{2}(a+b)$,

$$\mathbf{E}[(X - \tfrac{1}{2}(a+b))\,f'(X)] = \tfrac{1}{2}(f(a)+f(b)) - \tfrac{1}{b-a}\int_a^b f(t)\,\mathrm{d}t.$$

Hence, easy to write the details and get exactly Hadamard's $(*)$.

Play a Lottery: you may win one or two million dollars?!

Standard die: $\{1, 2, 3, 4, 5, 6\}$, each with probability $\frac{1}{6}$.

House/Casino Game Rules: You can play with n dice, where

$$n = 1, 2, 3, 4, \mathbf{5}, 6, 7, 8, 9, \mathbf{10}, 11, 12, 13, 14, \mathbf{15}, 16, 17, 18, 19, \mathbf{20}.$$

Toss your dice, get points: $X_1, \ldots, X_n$. Find the **sum** $\Sigma_n = X_1 + \ldots + X_n$, and the **product** $\Pi_n = X_1 \cdots X_n$. You win only if $\Sigma_n = \Pi_n$.

Play with **black** n, pay $\$n$ and if $\Sigma_n = \Pi_n$, win $\$\, n^3$. With **red** n:

n=5, pay $100, win $100,000;

n=10, pay $200, win $200,000;

n=15, pay $1,000, win $1,000,000;

n=20, pay $2,000, win $2,000,000.

Question: Are you ready to play? Why are 5, 10, 15 and 20 so special?

Reformulate the Problem: Start with a fixed set $M_6 = \{1, 2, 3, 4, 5, 6\}$ and let $n \in \mathbb{N}$. Choose n numbers $x_1, \dots, x_n$ from M_6, independently (possible repetitions) and solve the following **Diophantine equation**:

$$x_1 + \ldots + x_n = x_1 \cdots x_n. \tag{3}$$

Comment: (3) has a solution $\Leftrightarrow$ $W_n = \mathbf{P}[\Sigma_n = \Pi_n] > 0$, chance to win. If no solution to (3), your probability is $W_n = 0$, you win nothing!

Case 1: $n = 1 \Rightarrow$ trivial, $W_1 = 1$.

Case 2: $n = 2 \Rightarrow (2, 2) \Rightarrow W_2 = \frac{2!}{2!} \frac{1}{6^2}$.

Case 3: $n = 3 \Rightarrow (1, 2, 3) \Rightarrow W_3 = \frac{3!}{1!1!1!} \frac{1}{6^3}$.

Case 4: $n = 4 \Rightarrow (1, 1, 2, 4) \Rightarrow W_4 = \frac{4!}{2!1!1!} \frac{1}{6^4}$.

Case 5: $\mathbf{n = 5} \Rightarrow (1, 1, 1, 2, 5)$, $(1, 1, 1, 3, 3)$, and $(1, 1, 2, 2, 2) \Rightarrow$

$$W_5 = \left(\frac{5!}{3!1!1!} + \frac{5!}{3!2!} + \frac{5!}{2!3!}\right) \frac{1}{6^5}.$$

Exercise: Solve (3), see $W_n > 0$ for $n = 6, 7, 8, 9, \mathbf{10}, 11, 12, 13, 14$.

Surprising Fact: For $\mathbf{n = 15}$, eq. (3) does not have a solution!
I.e. $\Sigma_{15} = \Pi_{15}$ never happens, no win with 15 dice! [Sept'13, real story!]

Next: $W_n > 0$ for $n = 16, 17, 18, 19$, but for $\mathbf{n = 20}$, again $W_{20} = 0$.

Interestingly, $W_{21} > 0$, $W_{22} = 0$, $W_{23} > 0$, $W_{24} = 0$ followed by $W_n > 0$ for $n = 25, \ldots, 30$, and then for $n = 31$, again, $W_{31} = 0$, etc.

Def: Casino number is n for which eq. (3) does not have a solution; otherwise, n is a **usual number**.
Casino numbers: $38, 49, 95, 255, 529, 983$. They appear very irregularly.

Recent Result: There are infinitely many casino numbers and infinitely many usual numbers.

Question: We have found that $W_5 > 0$, $W_{10} > 0$, $W_{15} = 0$, $W_{20} = 0$.
Good reason for **15, 20** to be red. **Why are 5 and 10 red?**

General Model: Boxes, in each m balls, $m \in \mathbb{N}$, the numbers from $M_m = \{1, 2, \ldots, m\}$ used to mark the balls, $1 \leftrightarrow 1$, *standard numbering*.

Plato perfect solids: $m = 4$ for **tetrahedron**, $m = 6$ for **cube**, $m = 8$ for **octahedron**, $m = 12$ for **dodecahedron**, and $m = 20$ for **icosahedron**.

Choose n boxes, select randomly one ball from each, get $x_1, x_2, \ldots, x_n$. n is called a **casino number** if there is no solution to the Diophantine eq.

$$x_1 + x_2 + \ldots + x_n = x_1 x_2 \cdots x_n \quad \text{in the set} \quad M_m = \{1, 2, \ldots, m\}.$$

Otherwise, n is called a **usual number**.

Statement: Take arbitrary $m \in \mathbb{N}$ and let $M_m = \{1, 2, \ldots, m\}$.

(a) For any n we can tell exactly, this n is, or is not, a casino number.

(b) Both the casino numbers and the usual numbers are infinitely many.

Remark: With special, *nonstandard numbering* of the balls in the boxes, we can achieve: (i) all n are casino numbers; (ii) all n are usual numbers.

Corollary: For any m there is a minimal casino number, n_m. Examples:

$m = 4$ (tetrahedrons) $n_4 = 6$, next is 9

$m = 6$ (cubes) $n_6 = 15$, next is 20

$m = 8$ (octahedrons) $n_8 = 24$, next is 34

$m = 12$ (dodecahedrons) $n_{12} = 24$, next is 44

$m = 20$ (icosahedrons) $n_{20} = 24$, next is 80

Comment: For fixed m, M_m and T, a program can be written to produce all casino numbers in $[1, T]$. For $m = 6$, $M_6 = \{1, 2, 3, 4, 5, 6\}$ and $T = 100,000$, there are 98,417 casino numbers and 1,583 usual numbers.

Result: If p_T is the 'density' of the casino numbers in the interval $[1, T]$, then $p_T \to 1$ as $T \to \infty$. There is a T_0 such that with probab $\sim \frac{1}{2}$ a randomly chosen n from $[1, T_0]$ is a casino number. Similarly, for any p.

Mixed game: Choose n and play with any combination of n Plato solids.

200 years, K. Weierstrass **Th (1885)** Any continuous function on a closed bounded interval can be approximated uniformly by polynomials.

Bernstein Th (1912): Let $f(x),\ x \in [0,1]$ be a continuous function,

$$B_n(x) = \sum_{k=0}^{n} f\left(\frac{k}{n}\right)\binom{n}{k}x^k(1-x)^{n-k} \quad (\textbf{Bernstein Polynomials}).$$

Then, $B_n(x) \to f(x)$ uniformly on $[0,1]$; (sup norm) $\omega(f, n^{-1/2})$.

Proof: >100 **years** ago! Chebyshev Ineq. $\mathbf{P}[|X-a|>\varepsilon] \le \frac{1}{\varepsilon^2} Var[X]$.

Idea: n Bernoulli trials, $\perp\!\!\!\perp$, $X_i = 1$ or 0 with probab x and $1-x$,

$S_n = X_1 + \ldots + X_n \Rightarrow S_n \sim Bin(n,x)$, $\mathbf{P}[S_n = k] = \binom{n}{k}x^k(1-x)^{n-k}$.

$$\frac{S_n}{n} \to x, \quad n \to \infty \ (\textbf{Bernoulli LLN}) \ \Rightarrow \ B_n(x) = \mathbf{E}\left[f\left(\frac{S_n}{n}\right)\right] \to f(x).$$

In 2013: **300 years of "Ars Conjectandi", by Jacob Bernoulli.**

Uspensky Problem (AMM, 1932): Show that

$$\int_0^1 \cdots \int_0^1 \frac{x_1^2 + \cdots + x_n^2}{x_1 + \cdots + x_n} \mathrm{d}x_1 \cdots \mathrm{d}x_n \to \frac{2}{3} \quad \text{as} \quad n \to \infty.$$

Solutions rely essentially on the powers 2 and 1.

Next, with π and e the known classical constants, show that

$$\int_0^1 \cdots \int_0^1 \frac{x_1^\pi + \cdots + x_n^\pi}{x_1^\mathrm{e} + \cdots + x_n^\mathrm{e}} \mathrm{d}x_1 \cdots \mathrm{d}x_n \to \frac{\mathrm{e}+1}{\pi+1} \quad \text{as} \quad n \to \infty.$$

General result: For functions φ, f and g, all > 0, $\sim$ integr., as $n \to \infty$,

$$\frac{1}{c_0^n} \int \cdots \int \varphi(x_1) \cdots \varphi(x_n) \frac{f(x_1) + \cdots + f(x_n)}{g(x_1) + \cdots + g(x_n)} \mathrm{d}x_1 \cdots \mathrm{d}x_n \to \frac{\int f(x)\varphi(x)\mathrm{d}x}{\int g(x)\varphi(x)\mathrm{d}x}$$

Hint: $X_n \to X \Rightarrow \mathbf{E}[H(X_n)] \to \mathbf{E}[H(X)]$, H cont. bound. If $X = c$, LLN.

Riemann Zeta Function: $\zeta(z) = \sum_{n=1}^{\infty} \frac{1}{n^z}$, $\operatorname{Re} z > 1$.

Basel problem: Find $\zeta(2) = \sum_{n=1}^{\infty} \frac{1}{n^2} =?$ **Euler (1735):** $\zeta(2) = \frac{\pi^2}{6}$.

New Probabilistic Proof: We need a r.v. $X \sim HS$, hyperbolic secant,

$$f(x) = \frac{1}{\pi} \frac{1}{\cosh x} = \frac{2}{\pi} \frac{1}{e^x + e^{-x}}, \ x \in \mathbb{R}.$$

Next, take X_1, X_2, indep. as X, find density f_2 of $X_1 + X_2$, convolution

$$f_2(x) = \frac{4}{\pi^2} \frac{x}{e^x - e^{-x}}, \ x \in \mathbb{R} \ \text{ and use that } \int_{-\infty}^{\infty} f_2(x)\, dx = 1 :$$

$1 = \frac{8}{\pi^2} \int_0^{\infty} \frac{xe^{-x}}{1-e^{-2x}} dx = \frac{8}{\pi^2} \int_0^{\infty} xe^{-x} \sum_{k=0}^{\infty} e^{-2kx} dx = \frac{8}{\pi^2} \sum_{k=0}^{\infty} \frac{1}{(2k+1)^2}$.

$\Rightarrow \sum_{k=0}^{\infty} \frac{1}{(2k+1)^2} = \frac{\pi^2}{8}$. But $\zeta(2) = \sum_{odd} + \sum_{even} = \sum_{k=0}^{\infty} \frac{1}{(2k+1)^2} + \frac{1}{4} \zeta(2)$.

Infinite Systems of Functional Equations: $f = ?$

Statement 1c: Each of the following systems

$$\int_0^\infty x^k\, f_1(x)\,\mathrm{d}x = k! \quad \text{and} \quad \int_0^\infty x^k\, f_2(x)\,\mathrm{d}x = (2k)!, \ \ k = 1, 2, \ldots$$

has a unique solution: $f_1(x) = \mathrm{e}^{-x}, \ f_2(x) = \frac{1}{2}x^{-1/2}\mathrm{e}^{-x^{1/2}}, \ x > 0.$

Statement 2c: The following system

$$\int_0^\infty x^k\, f_3(x)\,\mathrm{d}x = (3k)!, \ \ k = 1, 2, \ldots$$

has infinitely many solutions, namely:

$$f_3(x) = \frac{1}{3}\,x^{-2/3}\,\mathrm{e}^{-x^{1/3}}\left(1 + \varepsilon\,\sin(\sqrt{3}\,x^{1/3} - \pi/3)\right), \ \ x > 0, \ \varepsilon \in [-1, 1].$$

Infinite Systems with Infinitely Many Unknowns:

Unknown are infinitely many real numbers $x_1 < x_2 < \ldots < x_n < \ldots$ and $p_1, p_2, \ldots, p_n, \ldots$, where all $p_n > 0$ and $p_1 + p_2 + \cdots + p_n + \cdots = 1$.

Statement 1d: There is no solution to each of the systems:

$$\sum_{n=1}^{\infty} x_n^k\, p_n = k!,\ \ k = 1, 2, \ldots \quad \text{and} \quad \sum_{n=1}^{\infty} x_n^k\, p_n = (2k)!,\ \ k = 1, 2, \ldots.$$

Statement 2d: There are infinitely many solutions to the system

$$\sum_{n=1}^{\infty} x_n^k\, p_n = (3k)!,\ \ k = 1, 2, \ldots.$$

Question: Idea? What is behind? ('c' = continuous, 'd' = discrete)

Idea: r.v. $\xi \sim Exp(1)$, $e^{-x}, x > 0$, $X = \xi^2$, $Y = \xi^3$, finite moments:
$m_k(\xi) = \mathbf{E}[\xi^k] = k!$, $m_k(X) = \mathbf{E}[X^k] = (2k)!$, $m_k(Y) = \mathbf{E}[Y^k] = (3k)!$

Fact 1: $\xi \sim Exp(1)$, is the only distr with the moments $\{k!\}$. Next, $X = \xi^2$, density $\frac{1}{2}x^{-1/2}e^{-x^{1/2}}$, the only distr with the moments $\{(2k)!\}$. Therefore, no other distributions $\Rightarrow$ Statements 1c, 1d.

Fact 2: $Y = \xi^3$, density $f_3(x) = \frac{1}{3}x^{-2/3}e^{-x^{1/3}}$, the distrib. is non-unique.

Berg & co.: There are infinitely many continuous and infinitely many discrete distributions all having moments $\{(3k)!\}$ $\Rightarrow$ Statement 2c, 2d.

Open Question: Find at least one solution to the system in the above Statement 2d.

Two Exercises and Another Open Question:

1. Prove that

$$\frac{1}{(\sin\frac{\pi}{5})^2}+\frac{1}{(\sin\frac{2\pi}{5})^2}=4; \qquad \frac{1}{(\sin\frac{\pi}{7})^2}+\frac{1}{(\sin\frac{2\pi}{7})^2}+\frac{1}{(\sin\frac{3\pi}{7})^2}=8.$$

2. Prove that

$$\frac{1}{1\cdot 2\cdot 3}+\frac{1}{4\cdot 5\cdot 6}+\frac{1}{7\cdot 8\cdot 9}+\cdots=\frac{\pi\sqrt{3}}{12}-\frac{1}{4}\log 3; \quad \sum_{n=1}^{\infty}\frac{1}{\binom{2n}{n}}=\frac{9+2\sqrt{3}\pi}{27}.$$

Open Question: Prove or disprove that **2, 4, 8, 64, 2048** are the only powers of 2 which are **written by using only even digits**. Any other power of 2 needs at least one odd digit.

Final Comments:

• **Kolmogorov-Feynman-Kac formulas** for solving 2nd order elliptic or parabolic PDEs. The solutions can be represented as expectations of functionals of Markov processes from Itô type stochastic differential equations. Hence, need of Stochastic Calculus. Applications: MCMC ...

• **Probabilistic methods in Number Theory and Graph Theory**
Work by Erdös, Spencer, Kubilius, Siegel, Bollobas, ...

• **Material Appropriate for:**
Mini-Courses, Full-Time Courses, Research Projects for Students.
And, perhaps, Challenges for Professionals!

References

Aigner & Ziegler: Erdös *Proofs from THE BOOK*

Alon & Spencer: *Probabilistic Method*

Bass: *Probabilistic Techniques in Analysis*

Billingsley: *Probability and Measure*

Bollobas: *Random Graphs*

Durrett: *Brownian Motion and Martingales in Analysis*

Feller: *An Introduction to the Theory of Probability* **2**

Freidlin: *Markov Processes and Differential Equations*

Sachkov: *Probabilistic Methods in Combinatorial Analysis*

Shiryaev: *Probability*

Papers by: M Yor, A DasGupta, J Galambos, J-F Le Gall, M Tamaki