Föreläsning 10: Diskret sannolikhetsteori, fortsättning · 1MA020

Vilhelm Agdur[1]

[1] vilhelm.agdur@math.uu.se

22 februari 2023

Vi fortsätter att diskutera diskret sannolikhetsteori, och introducerar slumpvariabler och deras väntevärden.

Vi använder den teori vi byggt upp för att bevisa några fler resultat inom kombinatoriken.

Slumpvariabler

Hittills är vad vi har sett bara hälften av vad man intuitivt tänker ingår i sannolikhetsteorin – vi har diskuterat slumpmässiga *händelser*, som antingen inträffar eller inte, men vi har inte definierat slumpmässiga tal. Frågan om ifall det kommer att regna imorgon eller inte kan vi modellera i vår formalism, men inte frågan om hur många millimeter det kommer regna.

Definition 1. Givet ett sannolikhetsrum (Ω, μ) är en *slumpvariabel* X som tar värden i V en funktion $X : \Omega \to V$. Givet varje utfall tar alltså vår slumpvariabel ett visst värde, och givet varje[2] delmängd $A \subseteq V$ blir $X \in A$ en händelse – specifikt är det händelsen

$$\{\omega \in \Omega \mid X(\omega) \in A\} = X^{-1}(A).$$

[2] Detta är lite av en lögn i det allmänna fallet, eftersom det kan finnas *väldigt* skumma delmängder till V, men så länge vi tänker oss våra diskreta sannolikhetsrum är det sant.

Det allra vanligaste fallet är när $V = \mathbb{R}$ eller någon delmängd till $\mathbb{R}$. I många introtexter om sannolikhetsteori *definierar* man att slumpvariabler tar värden i $\mathbb{R}$ – men eftersom vi sysslar med kombinatorik kommer vi att vilja ha mer exotiska slumpvariabler, som slumpmässiga permutationer eller slumpmässiga mängder.

Exempel 2. Låt oss återbesöka vårt exempel med ett tärningskast. Vi konstaterade att vi kan ta $\Omega = \{1, 2, 3, 4, 5, 6\}$ och $\mu(\omega) = 1/6$ för alla $\omega \in \Omega$.

Vi kan naturligt betrakta vårt tärningskast som en slumpvariabel – i detta fall blir det en mycket enkel funktion, $X : \Omega \hookrightarrow \mathbb{R}$ skickar helt enkelt varje ω på sig självt.

Vårt tärningskast är ett specialfall av ett mer allmänt fenomen, som det kommer vara bekvämt att ha en terminologi för.

Definition 3. Givet en ändlig mängd V är ett *likformigt fördelat slumpmässigt element av* V en slumpvariabel X sådan att $\mathbb{P}(X = v) = \frac{1}{|V|}$ för varje $v \in V$.[3] Alla element av V är alltså lika sannolika. Vi kan skriva detta som

$$X \overset{u}{\in} V.$$

[3] Vill man göra detta fullständigt rigoröst i vår formalism kan man säga att X är definierad på sannolikhetsrummet (V, μ) där $\mu(v) = \frac{1}{|V|}$ för alla $v \in V$, och $X : V \to V$ är identitetsfunktionen.

Men det blir väldigt många abstrakta ord för att inte säga så mycket alls som vi inte redan sade när vi definierade X som att den blir lika med varje element i V med samma sannolikhet.

Om någon säger att "vi låter X vara en slumpmässig graf / träd / mängd / etc." utan att specificera hur X är fördelad menar de att den är likformig.

Vi vet att om vi slår vår tärning många gånger kommer vi i genomsnitt att få upp 3.5. Hur gör vi den intuitionen rigorös?

Definition 4. *Väntevärdet* av en slumpvariabel X som tar värden i $\mathbb{R}$ ges av

$$\mathbb{E}[X] = \sum_{x \in X(\Omega)} x\mathbb{P}(X = x).$$

Vi tar alltså summan över alla tänkbara värden x för X, multiplicerar x med sannolikheten att X faktiskt blir x, och summerar. I specialfallet där X bara tar värden $0, 1, 2, \ldots$ blir alltså formeln

$$\mathbb{E}[X] = \sum_{k=0}^{\infty} k\mathbb{P}(X = k).$$

Exempel 5. Så om vi åter tar exemplet med tärningskastet så blir alltså väntevärdet

$$\begin{aligned}\mathbb{E}[X] &= 1\mathbb{P}(X = 1) + 2\mathbb{P}(X = 2) + \ldots + 6\mathbb{P}(X = 6) \\ &= \frac{1 + 2 + 3 + 4 + 5 + 6}{6} = \frac{7}{3} = 3.5\end{aligned}$$

precis som vi förväntade oss.

Ibland är det mer användbart att skriva definitionen av väntevärde på en alternativ form:

Lemma 6. *Det gäller att*

$$\mathbb{E}[X] = \sum_{\omega \in \Omega} X(\omega)\mu(\omega).$$

Bevis. Vi kan skriva

$$\begin{aligned}\mathbb{E}[X] &= \sum_{x \in X(\Omega)} x\mathbb{P}(X = x) \\ &= \sum_{x \in X(\Omega)} x \left(\sum_{\omega \in \Omega : X(\omega) = x} \mu(\omega) \right) \\ &= \sum_{x \in X(\Omega)} \sum_{\omega \in \Omega : X(\omega) = x} x\mu(\omega) \\ &= \sum_{x \in X(\Omega)} \sum_{\omega \in \Omega : X(\omega) = x} X(\omega)\mu(\omega) \\ &= \sum_{\omega \in \Omega} X(\omega)\mu(\omega).\end{aligned}$$

□

Eftersom vi definierat slumpvariabler som att de helt enkelt är funktioner från Ω kan vi göra all den algebra vi vanligen kan på funktioner in i $\mathbb{R}$. Till exempel är det, givet två slumpvariabler X och Y, helt väldefinierat att skriva $X + Y$, och det betyder precis vad vi förväntar oss att det skall betyda – vi slumpar ett X och ett Y och sedan adderar vi dem med varandra.

När vi nu har introducerat addition av slumpvariabler så kan vi bevisa vad som, i min mening, är en av de allra mest användbara satserna i hela matematiken.[4]

[4] Jag är så klart oerhört partisk, eftersom just gränslandet mellan kombinatorik och sannolikhetsteori är mitt område – men det är onekligen ett otroligt användbart resultat.

Lemma 7 (Väntevärdets linjäritet)**.** *Givet två slumpvariabler X och Y och två reella tal a och b gäller det att*

$$\mathbb{E}\left[aX + bY\right] = a\mathbb{E}\left[X\right] + b\mathbb{E}\left[Y\right].$$

Väntevärdet är alltså linjärt, som funktion från rummet av slumpvariabler in i $\mathbb{R}$.[5]

[5] Detta sätt att formulera det skrapar lite på ytan av en väldigt djup teori – väntevärden är nämligen "bara" integraler mot sannolikhetsmått, och samlingen av funktioner från Ω in i $\mathbb{R}$ blir ju ett vektorrum. Vi kan ge det vektorrummet en inre produkt genom att skriva $\langle X, Y\rangle = \mathbb{E}\left[XY\right]$, och vi har börjat med funktionalanalys.

Men detta är ju en kurs i kombinatorik, så att utforska detta får vänta till en framtida kurs för er.

Bevis. Vi använder den alternativa formeln för väntevärde vi fann i Lemma 6 och skriver

$$\begin{aligned}
\mathbb{E}\left[aX + bY\right] &= \sum_{\omega\in\Omega} (aX + bY)(\omega)\mu(\omega) \\
&= \sum_{\omega\in\Omega} (aX(\omega) + bY(\omega))\mu(\omega) \\
&= a\sum_{\omega\in\Omega} X(\omega)\mu(\omega) + b\sum_{\omega\in\Omega} Y(\omega)\mu(\omega) \\
&= a\mathbb{E}\left[X\right] + b\mathbb{E}\left[Y\right].
\end{aligned}$$

□

Övningar