Discrete Probability, Random Variables, and Bounds

Spørgsmål 3 fra Exam Questions

Kevin Vinther

December 10, 2023

Table of Contents i

An Introduction to Discrete Probability

- **Experiment**: En procedure som giver et resultat ud af et givet sæt a mulige resultater.
- **Sample Space**: Sættet af alle mulige resultater.
- **Event**: Subset af sample space.

Definition 1

Definition

If S is a finite nonempty sample space of equally likely outcomes, and E is an event, that is, a subset of S, then the *probability* of E is $p(E) = \frac{|E|}{|S|}$.

- *Implikationer:*
- Givet, at E er en event, og S sample space:
- $0 \leq |E| \leq |S|$, fordi $E \subseteq S$. Dermed $0 \leq p(E) = |E|/|S| \leq 1$.

Example 1

- En boks indeholder fire blå bolde og fem røde. Hvad er sandsynligheden at en bold der er valgt tilfældigt er blå?

Example 1

- En boks indeholder fire blå bolde og fem røde. Hvad er sandsynligheden at en bold der er valgt tilfældigt er blå?
- $\frac{4}{9}$

Example 2

- Hvad er sandsynligheden for at, når 2 terninger bliver rullet, at summen af de to tal på terningen er 7?
- $6^2 = 36$ mulige udfald. (Af product rule).

Example 2

- Hvad er sandsynligheden for at, når 2 terninger bliver rullet, at summen af de to tal på terningen er 7?
- $6^2 = 36$ mulige udfald. (Af product rule).
- Der er 6 udfald der ender i 7: (1,6), (2,5), (3,4), (4,3), (5,2), og (6,1)

Example 2

- Hvad er sandsynligheden for at, når 2 terninger bliver rullet, at summen af de to tal på terningen er 7?
- $6^2 = 36$ mulige udfald. (Af product rule).
- Der er 6 udfald der ender i 7: (1,6), (2,5), (3,4), (4,3), (5,2), og (6,1)
- Dermed: $\frac{6}{36} = \frac{1}{6}$

Eksempel 5

- Hvad er sandsynligheden for at en hånd i poker har 4 kort af én slags (samme tal)? (Hint: Product rule)

Eksempel 5

- Hvad er sandsynligheden for at en hånd i poker har 4 kort af én slags (samme tal)? (Hint: Product rule)
- Antallet af hænder med 5 kort med fire af én slags er produktet af antallet måder man kan vælge én slags på, antallet af måder du kan vælge disse 4 af slagsen ud af de fire i dækket, og antallet af måder du kan vælge det femte kort. Det er:

Eksempel 5

- Hvad er sandsynligheden for at en hånd i poker har 4 kort af én slags (samme tal)? (Hint: Product rule)
- Antallet af hænder med 5 kort med fire af én slags er produktet af antallet måder man kan vælge én slags på, antallet af måder du kan vælge disse 4 af slagsen ud af de fire i dækket, og antallet af måder du kan vælge det femte kort. Det er:
- $\binom{13}{1}\binom{4}{4}\binom{48}{1}$. Næste: Hvor mange forskellige hænder af 5 kort er der?

Eksempel 5

- Hvad er sandsynligheden for at en hånd i poker har 4 kort af én slags (samme tal)? (Hint: Product rule)
- Antallet af hænder med 5 kort med fire af én slags er produktet af antallet måder man kan vælge én slags på, antallet af måder du kan vælge disse 4 af slagsen ud af de fire i dækket, og antallet af måder du kan vælge det femte kort. Det er:
- $\binom{13}{1}\binom{4}{4}\binom{48}{1}$. Næste: Hvor mange forskellige hænder af 5 kort er der?
- $\binom{52}{5}$. Næste: Find sandsynligheden

Eksempel 5

- Hvad er sandsynligheden for at en hånd i poker har 4 kort af én slags (samme tal)? (Hint: Product rule)
- Antallet af hænder med 5 kort med fire af én slags er produktet af antallet måder man kan vælge én slags på, antallet af måder du kan vælge disse 4 af slagsen ud af de fire i dækket, og antallet af måder du kan vælge det femte kort. Det er:
- $\binom{13}{1}\binom{4}{4}\binom{48}{1}$. Næste: Hvor mange forskellige hænder af 5 kort er der?
- $\binom{52}{5}$. Næste: Find sandsynligheden
- $\frac{\binom{13}{1}\binom{4}{4}\binom{48}{1}}{\binom{52}{5}} = \frac{13 \cdot 1 \cdot 48}{2598960} \approx 0.00024$

Eksempel 7 i

- Hvad er sandsynligheden at tallene 11, 4, 17, 39 og 23 bliver taget i den orden fra en spand med 50 bolde med tallene 1..50, hvis (a) bolden der er valgt ikke bliver lagt tilbage i spanden før den næste bold blive valgt og (b) bolden bliver lagt tilbage før den næste bold bliver valgt?
- Lad os starte med spørgsmål (a) (hint: Hvor mange måder kan man vælge boldene på?):

Eksempel 7 i

- Hvad er sandsynligheden at tallene 11, 4, 17, 39 og 23 bliver taget i den orden fra en spand med 50 bolde med tallene 1..50, hvis (a) bolden der er valgt ikke bliver lagt tilbage i spanden før den næste bold blive valgt og (b) bolden bliver lagt tilbage før den næste bold bliver valgt?
- Lad os starte med spørgsmål (a) (hint: Hvor mange måder kan man vælge boldene på?):
- Man kan vælge 5 bolde på $50 \cdot 49 \cdot 48 \cdot 47 \cdot 46 = 254251200$ måder. Hvad er sandsynligheden så for at vælge præcis den orden der er beskrevet tidligere?

Eksempel 7 i

- Hvad er sandsynligheden at tallene 11, 4, 17, 39 og 23 bliver taget i den orden fra en spand med 50 bolde med tallene 1..50, hvis (a) bolden der er valgt ikke bliver lagt tilbage i spanden før den næste bold blive valgt og (b) bolden bliver lagt tilbage før den næste bold bliver valgt?
- Lad os starte med spørgsmål (a) (hint: Hvor mange måder kan man vælge boldene på?):
- Man kan vælge 5 bolde på $50 \cdot 49 \cdot 48 \cdot 47 \cdot 46 = 254251200$ måder. Hvad er sandsynligheden så for at vælge præcis den orden der er beskrevet tidligere?
- $\frac{1}{254251200}$.
- Bogen kalder dette **sampling without replacement**.

Eksempel 7 ii

- Hvad med (b)? Hvor mange måder kan man vælge 5 bolde på så?

Eksempel 7 ii

- Hvad med (b)? Hvor mange måder kan man vælge 5 bolde på så?
- 50^5. Husk at antallet af permutationer med gentagelser er n^r, og her er $n = 50, r = 5$.

Eksempel 7 ii

- Hvad med (b)? Hvor mange måder kan man vælge 5 bolde på så?
- 50^5. Husk at antallet af permutationer med gentagelser er n^r, og her er $n = 50, r = 5$.
- Måden man finder sandsynligheden på har ikke ændret sig her.
- $\frac{1}{312500000}$
- Bogen kalder dette **sampling with replacement**.
- Hvad forskellen er mellem det og permutations with repetition er et godt spørgsmål.

Theorem 1

Tid til første theorem!

Theorem (Theorem 1)

Let E be an event in a sample space S. The probability of the event $\overline{E} = S - E$, the complementary event of E is given by

$$P(\overline{E}) = 1 - p(E)$$

Theorem 1 Bevis

Theorem 1 Bevis.

For at finde sandsynligheden af begivenheden $\overline{E} = S - E$, notér at $|\overline{E}| = |S| - |E|$. Dermed:

$$p(\overline{E}) = \frac{|S| - |E|}{|S|} = 1 - \frac{|E|}{|S|} = 1 - p(E)$$

$\square$

Eksempel 8

- En sekvens af 10 bits er tilfældigt genereret. Hvad er sandsynligheden for at mindst en af disse bits er 0?
- Hint: Brug complement.

Eksempel 8

- En sekvens af 10 bits er tilfældigt genereret. Hvad er sandsynligheden for at mindst en af disse bits er 0?
- Hint: Brug complement.
- Lad E være tilfældet at mindst en af de 10 bits er 0. Så er $\overline{E}$ sandsynligheden for at alle bitsne er 1.

Eksempel 8

- En sekvens af 10 bits er tilfældigt genereret. Hvad er sandsynligheden for at mindst en af disse bits er 0?
- Hint: Brug complement.
- Lad E være tilfældet at mindst en af de 10 bits er 0. Så er $\overline{E}$ sandsynligheden for at alle bitsne er 1.
- $P(E) = 1 - p(\overline{E}) = 1 - \frac{|\overline{E}|}{|S|} = 1 - \frac{1}{2^{10}} = 1 - \frac{1}{1024} = \frac{1023}{1024}$

Theorem 2

Theorem

Let E_1 and E_2 be events in the sample space S. Then

$$p(E_1 \cup E_2) = p(E_1) + P(E_2) - p(E_1 \cap E_2).$$

Theorem 2 Bevis

Theorem 2 Bevis.

Vi bruger subtraction rule for sæt (inclusion-exclusion for 2 sæt):

$$|E_1 \cup E_2| = |E_1| + |E_2| - |E_1 \cap E_2|$$

Dermed:

$$\begin{aligned}
p(E_1 \cup E_2) &= \frac{|E_1 \cup E_2|}{|S|} \\
&= \frac{|E_1| + |E_2 - |E_1 \cap E_2|}{|S|} \\
&= \frac{|E_1|}{|S|} + \frac{|E_2|}{|S|} - \frac{|E_1 \cap E_2|}{|S|} \\
&= p(E_1) + p(E_2) - p(E_1 \cap E_2)
\end{aligned}$$

$\square$

Eksempel 9

- Hvad er sandsynligheden for at et positivt heltal valgt tilfældligt fra et sæt af positive heltal der ikke overgår 100 er deleligt med 2 eller 5?
- Hint: Definér E_1 og E_2 som værende tal der er delelige med hhv. 2 og 5.

Eksempel 9

- Hvad er sandsynligheden for at et positivt heltal valgt tilfældligt fra et sæt af positive heltal der ikke overgår 100 er deleligt med 2 eller 5?
- Hint: Definér E_1 og E_2 som værende tal der er delelige med hhv. 2 og 5.
- $E_1 =$ tal der er delelige med 2, $E_2 =$ tal der er delelige med 5.

Eksempel 9

- Hvad er sandsynligheden for at et positivt heltal valgt tilfældligt fra et sæt af positive heltal der ikke overgår 100 er deleligt med 2 eller 5?
- Hint: Definér E_1 og E_2 som værende tal der er delelige med hhv. 2 og 5.
- $E_1 =$ tal der er delelige med 2, $E_2 =$ tal der er delelige med 5.
- $|E_1| = 50$, $|E_2| = 20$, $|E_1 \cup E_2| = 10$.
- Vi ved at $p(E_1 \cup E_2) = p(E_1) + P(E_2) - p(E_1 \cap E_2)$. Så: $= \frac{50}{100} + \frac{20}{100} - \frac{10}{100} = \frac{3}{5}$

Probabilistic Reasoning

- Probabilistic Reasoning er at bruge sandsynlighedsregning til at ræsonnere sig frem til resultater. Et eksempel på hvor man kan gøre dette, er Monty Hall Three-Door Puzzle, som vi kommer til nu!

Eksempel 10: The Monty Hall Three-Door Puzzle

- Antag at du er en spiller i et game show.
- Du bliver bedt om at vælge en af tre døre.
- Den store pris er bag en af de tre døre, og de to andre døre leder til tab.
- Når du vælger en dør, vil værten lade dig enten beholde døren, eller vælge en ny. Du får ikke lov til at se hvad der er i døren hvis du vælger en ny.
- Vælger du en ny?

Eksempel 10 Løsning

- Ja du gør!
- Sandsynligheden for at du har valgt den rigtige dør først er $\frac{1}{3}$
- Sandsynligheden for at du har valg **forkert** er $\frac{2}{3}$.
- Hvis du valgte forkert, åbner game show værten en dør hvor der ingen præmie er bag.
- Dermed, kommer du **altid** til at vinde, hvis dit første valg var forkert, og du så vælger at skifte døre.
- Så, ved at skifte døre er chancen for at vinde $\frac{2}{3}$

Reflektion

- Hvad er et eksperiment?
- Sample space?
- Event?
- Laplace's Definition?
- Hvad er værdimængden for $P(E)$? (Givet at E er en event)
- Hvad er sampling with og without replacement?
- Hvordan fungerede beviset til $P(\overline{E})$?

Probability Theory

Assigning probabilities

We giver sandsynligheden $p(s)$ til hvert udfald s. We kræver at to betingelser bliver overholdt:

- $0 \leq p(s) \leq 1$for each$s \in S$ og
- $\sum_{s \in S} p(s) = 1$

Funktionen p bliver kaldt **probability distribution**.

Definition 1

Definition

Suppose that S is a set with n elements. The *uniform distribution* assigns the probability $1/n$ to each element of S.

- Basically, uniform distribution er hvor hvert udfald har lige stor sandsynlighed for at ske (og det er $1/n$ fordi vi ved at $\sum_{s \in S} p(s) = 1$.)

Definition 2

Definition

The *probability* of the event E is the sum of the probabilities of the outcomes in E. That is

$$p(E) = \sum_{s \in E} p(s)$$

- Bare det jeg sagde på sidste slide, lol. Jeg er ikke helt sikker på hvorfor det er gentaget her.

Theorem 1

Theorem

If $E_1, E_2, \ldots$ is a sequence of pairwise disjoint events in a sample space S, then

$$p\left(\bigcup_i E_i\right) = \sum_i p(E_i)$$

- Beviset undlades, men jeg mindes at det bringes op i KT 13.1

Definition 3

Definition

Let E and F be events with $p(F) > 0$. The *conditional probability* of E given F, denoted by $p(E|F)$, is defined as

$$p(E|F) = \frac{p(E \cap F)}{p(F)}$$

Eksempel 3

- En bit streng af længde 4 er tilfældligt generet, således at hver af de 16 bits af længde 4 er lige sandsynlige. Hvad er sandsynligheden at den har mindst to 0er i træk (Til oplægsholder: næste slide har lidt hjælp)

Eksempel 3

- En bit streng af længde 4 er tilfældligt generet, således at hver af de 16 bits af længde 4 er lige sandsynlige. Hvad er sandsynligheden at den har mindst to 0er i træk (Til oplægsholder: næste slide har lidt hjælp)
- Lad E være begivenheden at en bit streng af længde 4 har mindst to 0er i træk.
- Lad F være begivenheden at den første bit af en streng af længde fire er et 0.
- Hvad er sandsynligheden så?

Eksempel 3

- En bit streng af længde 4 er tilfældligt generet, således at hver af de 16 bits af længde 4 er lige sandsynlige. Hvad er sandsynligheden at den har mindst to 0er i træk (Til oplægsholder: næste slide har lidt hjælp)
- Lad E være begivenheden at en bit streng af længde 4 har mindst to 0er i træk.
- Lad F være begivenheden at den første bit af en streng af længde fire er et 0.
- Hvad er sandsynligheden så?
- $p(E|F) = \frac{p(E \cap F)}{p(F)}$
- Siden $E \cap F = \{0000, 0001, 0010, 0011, 0100\}$, og dermed, $|E \cap F| = 5$, så må $p(E \cap F) = 5/16$. Ydermere er $p(F) = 1/2 = 8/16$. Derfor:
- $p(E|F) = \frac{5/16}{1/2} = \frac{5}{8}$

Independence

- En begivenhed (event) er independent (uafhængig), hvis tidligere udfald ikke har effekt på sandsynligheden af begivenheden.

Definition (Definition 4)

The events E and F are *independent* if and only if $p(E \cap F) = p(E)p(F)$

Eksempel 5

- Antag at E er begivenheden at en tilfældigt genereret big streng af længde 4 begynder med et 1, og F er begivenheden at denne bit streng har præcis et lige antal 1'er. Er E og F uafhængige, hvis de 16 bits er lige sandsynlige?

Eksempel 5

- Antag at E er begivenheden at en tilfældigt genereret big streng af længde 4 begynder med et 1, og F er begivenheden at denne bit streng har præcis et lige antal 1'er. Er E og F uafhængige, hvis de 16 bits er lige sandsynlige?
- $E = \{100, 1001, 1010, 1011, 1100, 1101, 1110, 1111\}$
- $F = \{0000, 0011, 0101, 0110, 1001, 1010, 1100, 1111\}$.

Eksempel 5

- Antag at E er begivenheden at en tilfældigt genereret big streng af længde 4 begynder med et 1, og F er begivenheden at denne bit streng har præcis et lige antal 1'er. Er E og F uafhængige, hvis de 16 bits er lige sandsynlige?
- $E = \{100, 1001, 1010, 1011, 1100, 1101, 1110, 1111\}$
- $F = \{0000, 0011, 0101, 0110, 1001, 1010, 1100, 1111\}$.
- Dermed: $p(E) = p(F) = 8/16 = 1/2$
- Fordi $E \cap F = \{1111, 1100, 1010, 1001\}$, ser vi at:
- $p(E \cap F) = 4/16 = 1/4 = (1/2)(1/2) = p(E)p(F)$
- Dermed er E og F uafhængige.

Pairwise and Mutual Independence

Definition (Definition 5)

The events $E_1, E_2, \ldots, E_n$ are *pairwise independent* if and only if $p(E_i \cap E_j) = p(E_i)p(E_j)$ for all pairs of integers i and j with $1 \leq i < j \leq n$. These events are *mutually independent* if $p(E_{i_1} \cap E_{i_1} \cap \cdots \cap E_{i_m}) = p(E_{i_1})p(E_{i_2}) \cdots p(E_{i_m})$ whenever $i_j, j = 1, 2, \ldots, m$ are integers with $1 \leq i_1 < i_2 < \cdots < i_m \leq n$ and $m \geq 2$.

- Forskellen er, i sin essens, at pairwise independence tillader at en begivenhed kan påvirke mere end én begivenheds sandsynlighed, bare ikke kun én.
- Derudover, kan sandsynligheder godt være mutually independent uden at være pairwise independent.

Bernoulli Trials

- En **Bernoulli Trial** er et ekseperiment hvor der kun er to mulige udfald.
- Generelt er disse udfald kaldet *success* eller *fejl* (failure).
- Givet at p er sandsynligheden for success, og q for fejl, må $p + q = 1$.
- Bernoulli Trials er mutually independent, hvis conditional probability af success på et givet trial er p, givet hvilket som helst information om udfaldet af andre *trials*.

Eksempel 8

- En mønt er biased, så sandsynligheden for plat er $2/3$. Hvad er sandsynligheden for at præcis fire plat kommer up når man kaster mønten syv gange?

Eksempel 8

- En mønt er biased, så sandsynligheden for plat er $2/3$. Hvad er sandsynligheden for at præcis fire plat kommer up når man kaster mønten syv gange?
- Der er $2^7 = 128$ forskellige udfald. Antallet af måder fire ud af de 7 møntkast kan være plat på er $\binom{7}{4}$.
- Sandsynligheden for det præcise udfald er $(2/3)^4(1/3)^3$. Dermed er sandsynligheden:

Eksempel 8

- En mønt er biased, så sandsynligheden for plat er $2/3$. Hvad er sandsynligheden for at præcis fire plat kommer up når man kaster mønten syv gange?
- Der er $2^7 = 128$ forskellige udfald. Antallet af måder fire ud af de 7 møntkast kan være plat på er $\binom{7}{4}$.
- Sandsynligheden for det præcise udfald er $(2/3)^4(1/3)^3$. Dermed er sandsynligheden:
- $$\binom{7}{4}(2/3)^4(1/3)^3 = \frac{35 \cdot 16}{3^7} = \frac{560}{2187}$$

Theorem 2

Theorem (Theorem 2)

The probability of exactly k successes in n independent Bernoulli trials, with probability of success p and probability of failure $q = 1 - p$, is

$$\binom{n}{k} p^k q^{n-k}$$

Theorem 2 Bevis

Længe siden vi har set sådan et, hva?

Theorem 2 Bevis.

Når n Bernoulli trial bliver udført, er udfaldet en n-tuple $(t_1, t_2, \ldots, t_n)$, hvor $t_i = S$ (success) eller $t_i = F$ (fejl), for $i = 1, 2, \ldots, n$. Fordi de n trials er uafhængige, så er sandsynligheden af hvert udfald af de n trials af k succeser og $n - k$ fejl $p^k q^{n-k}$. Fordi der er $\binom{n}{k}$ $n-$tupler af S'er og F'er som har præcis k S'er, er sandsynligheden for præcis k successer:

$$\binom{n}{k} p^k q^{n-k}$$

□

Binomial Distribution

- Vi betegner $b(k; n, p)$ til at være sandsynligheden af k sucesser i n uafhængige Bernoulli trials med sandsynlighed for succes p og sandsynlighed for fejl $q = 1 - p$. Vi kalder denne funktion **binomial distribution**.

Random Variables

- Random Variables, kendt som at være et af de mest vildledende navne i matematik efter Dynamic Programmering, er ellers ret vigtigt i sandsynlighedsregning.
- De er dog ret simple, de kommer nok ikke til at fylde en specielt stor del af resten af kurset.

Definition 6

Definition

A *random variable* is a function from the sample space of an experiment to the set of real numbers. That is, a random variable assigns a real number to each posssible outcome.

- Så.. Det er bare en tilfældig variabel?
- Nej!
- Den er hverken tilfældig, og den er heller ikke variabel! Det er en funktion!!!

Definition 7

Definition

The *distribution* of a random variable X on a sample space S is the set of pairs $(r, p(X = r))$ for all $r \in X(S)$, where $p(X = r)$ is the probability that X takes the value r. (The set of pairs in this distribution is determined by the probabilities $p(X = r)$ for $r \in X(S)$.)

Eksempel 13: The Birthday Problem i

- (Dette er ikke en opgave lignende andre eksempler, jeg giver svaret, og forventer ikke at I kan det).
- Hvad er det minimum antal af folk der skal være i et rum således at sandsynligheden for at to af dem har samme fødselsdag er større end $1/2$?
- Vi antager at fødselsdagene er uafhængige.
- Vi antager at hver fødselsdag har lige sandsynlighed for at være præcis den dag, i.e. sandsynligheden for at fødselsdagen er på en præcis dag: $1/n$ hvor n er antallet af dage, som er:
- Vi antager at antallet af dage på et år er 366.
- Fremgangsmåde:

Eksempel 13: The Birthday Problem ii

- Vi må først udregne sandsynligheden p_n for at all n personer har forskellige fødselsdage
- Derefter taget vi $1 - p_n$, som så giver os sandsynligheden for at mindst to personer deler fødselsdag.

- Sandsynligheden for at to personer har samme fødselsdag er $\frac{365}{366}$. Dermed, ved at vi tager til j personer bliver det:
-
$$\frac{366(j-1)}{366} = \frac{367-j}{366}$$

Eksempel 13: The Birthday Problem iii

- Fordi vi har antaget at fødselsdagene af personerne i rummet er uafhængige, kan vi konkludere at sandsynligheden at de n personer i rummet har forskellige fødselsdage er:

$$p_n = \frac{365}{366}\frac{364}{366}\frac{363}{366}\cdots\frac{367-n}{366}$$

.

- Sandsynligheden for at to har samme fødselsdag er $1 - p_n$.
- Vi gider ikke til at bruge differentialregning eller whatever, så vi prøver os bare frem, og finder ud af at denne formel giver $n = 23, 1 - p_n \approx 0.506$.

Eksempel 14: Probability of a Collision in Hashing Functions i

- Denne her er god at kunne når vi kommer til hashing!
- Hvad er sandsynligheden for at ingen nøgler mapper til den samme lokation af én hashing funktion?
- Vi antager universal hashing function (i.e., sandsynligheden for at den rammer en specifik lokation er $1/m$, hvor m er antallet af slots.)
- Vi modellerer nøglerne til at være $k_1, k_2, \ldots, k_n$, hvor n er antallet af nøgler.
- Sandsynligheden for at $h(k_2) \neq h(k_1) = \frac{(m-1)}{m}$ fordi der er $m-1$ lokationer den kan komme til, ud af m i alt, og den første er allerede taget. Det fortsætter og bliver til:

Eksempel 14: Probability of a Collision in Hashing Functions ii

- $\frac{m-(j-1)}{m}$ hvor j er nøglenummeret (fra 1 opad). Dermed, givet at p_n er sandsynligheden for at $h(k_1) \neq h(k_2) \neq \cdots \neq h(k_n) = p_n$,

$$p_n = \frac{m-1}{m} \cdot \frac{m-2}{m} \cdot \cdots \cdot \frac{m-n+1}{m}$$

.

- Vi finder så sandsynligheden ved at tage komplement, i.e. $1 - p_n$.

Theorem 3

Theorem (Theorem 3)

The Probabilistic Method *If the probability that an element chosen at random from a S does not have a particular property is less than 1, there exists an elementi n S with this property.*

- Kan blive brugt til at lave nonkonstruktive eksistensbeviser (i.e., du viser at et element eksisterer, du viser bare ikke hvor det er, eller hvordan du finder det).

Theorem 4

Theorem

If k is an integer with $k \geq 2$, then $R(k, k) \geq 2^{k/2}$.

Vi ved at teoremet holder for $k = 2$ og $k = 3$, fordi $R(2,2) = 2, R(3,3) = 6$. (Dette var vist i kapitel 6.2).

Hvis vi så antager at $k \geq 4$, vil vi bruge den probabilistike metode til at vise om der er færre end $2^{k/2}$ personer til en fest, så er det muligt at ingen k af dem er *mutual friends or mutual enemies* (jeg gider ikke til at oversætte).

Vi antager at sandsynligheden for at to personer er venner eller fjender er lige stor.

Antag at der er n personer til festen. Dermed er der $\binom{n}{k}$ forskellige sæt af k personer til denne færst, hvilke vi skriver som $S_1, S_2, \ldots, S_{\binom{n}{k}}$ (altså, sættene af personer).

Theorem 4 Bevis ii

Lad E_i være begivenheden at alle k personer i S_i er enten *mutual friends* eller *mutual enemies*.

Sandsynligheden for at der enten er k *mutual friends* eller *mutual enemies* blandt de n personer er lig med

$$\bigcup_{i=1}^{\binom{n}{k}} E_i$$

.

Ifølge vores antagelse at det er lige sandsynligt for to personer at være fjender eller venner, så er sandsynligheden for begge $1/2$. Ydermere er der $\binom{k}{2} = \frac{k(k-1)}{2}$ par af personer i S_i, fordi der er k personer i S_i. Dermed er sandsynligheden for at alle k personer i S_i

Theorem 4 Bevis iii

er *mutual friends* og sandsynligheden for at alle k personer i S_i er *mutual enemies* er begge lig $(1/2)^{k(k-1)/2}$. Det følger at $p(E_i) = 2(1/2)^{k(k-1)/2}$.

Sandsynligheden for at der er enten k *mutual friends* eller k *mutual enemies* i gruppen af n personer er lig med $p(\bigcup_{i=1}^{\binom{n}{k}} E_i)$. Ved brug af Boole's Inequality, følger det at:

$$p\left(\bigcup_{i=1}^{\binom{n}{k}} \leq \sum_{i=1}^{\binom{n}{k}} p(E_i) = \binom{n}{k} \cdot 2\left(\frac{1}{2}\right)^{k(k-1)/2}\right)$$

Theorem 4 Bevis iv

Af exercise 21 i 6.4 har vi at $\binom{n}{k} \leq n^k/2^{k-1}$. Dermed,

$$\binom{n}{k} 2 \left(\frac{1}{2}\right)^{k(k-1)/2} \leq \frac{n^k}{2^{k-1}} 2 \left(\frac{1}{2}\right)^{k(k-1)/2}$$

Hvis $n < 2^{k/2}$, har vi at :

$$\frac{n^k}{2^{k-1}} 2 \left(\frac{1}{2}\right)^{k(k-1)/2} < \frac{2^{k(k/2)}}{2^{k-1}} 2 \left(\frac{1}{2}\right)^{k(k-1)/2} = 2^{2-(k/2} \leq 1$$

Reflektion i

- Hvad er probability distribution?
- Hvad er en uniform distribution?
- Hvad er conditional probability?
- Hvordan udregnes $P(E|F)$?
- Forklar, uden brug af matematik, hvad uafhængige begivenheder er.
- Forklar med matematik.
- Hvad er pairwise independence
- Hvad er mutual independence
- Hvad er forskellen?
- Ligheden?

Reflektion ii

- Hvad er en bernoulli trial?
- Hvad er en binomial distribution (binomial fordeling)?
- Hvad er en random variable?
- Er det en variable?
- Hvad er den probabilistiske metode?

Bayes' Theorem

Bayes' Theorem

- Bayes' Theorem er et teorem der kan hjælpe dig med at finde en sandsynlighed, baseret udelukkende på partial knowledge.

Eksempel 1 i

- Vi har to bokse. Den første indeholder to grønne bolde, og syv røde bolde. Den anden indeholder fire grønne bolde og tre grønne bolde.
- Bob vælger en bold ved først at vælge en af de to bokse tilfældigt, og derefter en bold tilfældeligt. Hvis bob har valgt en rød bold, hvad er sandsynligheden at han tog fra den første boks?
- Lad E være begivenheden at Bob har valgt en rød bold. $\overline{E}$ er begivenheden at Bob har valgt en grøn bold. Lad F være begivenheden at Bob har valgt fra den første boks ($\overline{F}$ er dermed begivenheden at han har valgt fra anden).

Eksempel 1 ii

- Vi vil gerne finde $P(F|E)$. Vi ved at $P(F|E) = \frac{P(F \cap E)}{P(E)}$. Kan vi bruge informationen givet til at finde både $P(F \cap E)$ og $P(E)$, så vi kan finde $P(F|E)$?
- For det første ved vi at vi kan finde sandsynligheden for at han vælger en rød bold, hvis han tager fra den første boks. $P(E|F) = 7/9$. Vi ved også sandsynligheden for at han tager en rød bold hvis han tager fra den anden boks: $P(E|\overline{F}) = 3/7$.
- Siden vi antager at Bob vælger en boks tilfældig, er $P(F) = P(\overline{F}) = 1/2$.

Eksempel 1 iii

- Vi ved at $p(E|F) = \frac{p(E \cap F)}{p(F)}$. Hvis vi ganger med $p(F)$ på begge sider får vi: $p(E|F) \cdot p(F) = \frac{p(E \cap F)}{p(F)} \cdot p(F) = p(E \cap F)$. Dermed kan vi finde $p(E \cap F)$ ved at have $p(E|F)$ og $p(F)$. Vi har etableret, at vi har denne information.
- $\frac{7}{9} \cdot \frac{1}{2} = \frac{7}{18}$
- Ligeledes: $p(E \cap \overline{F}) = p(E|\overline{F})p(\overline{F}) = \frac{3}{7} \cdot \frac{1}{2} = \frac{3}{14}$
- Da $E = (E \cap F) \cup (E \cap \overline{F})$ kan vi nu finde $p(E)$.
- $p(E) = p(E \cap F) + p(E \cap \overline{F}) = \frac{7}{18} + \frac{3}{14} = \frac{49}{126} + \frac{27}{126} = \frac{76}{126} = \frac{38}{63}$
- Dermed:
- $$p(F|E) = \frac{p(F \cap E)}{p(E)} = \frac{7/18}{38/63} = \frac{49}{76} \approx 0.645$$

Theorem 1

Theorem (Bayes' Theorem)

Suppose that E and F are events from a sample space S such that $p(E) \neq 0$ and $p(F) \neq 0$. Then

$$p(F|E) = \frac{p(E|F)p(F)}{p(E|F)p(F) + p(E|\overline{F})p(\overline{F})}$$

Definitionen af conditional probability fortæller os at $p(F|E) = p(E \cap F)/p(E)$ og $p(E|F) = p(E \cap F)/p(F)$. Derfor, $p(E \cap F) = p(F|E)p(E)$ og $p(E \cap F) = p(E|F)p(F)$. Hvis vi sætter disse lig hinanden, får vi:

$$p(F|E)p(E) = p(E|F)p(F)$$

Hvis vi dividerer begge disse sider med $p(E)$ får vi:

$$p(F|E) = \frac{p(E|F)p(F)}{p(E)}$$

Som det næste viser vi at $p(E) = p(E|F)p(F) + P(E|\overline{F})p(\overline{F})$. For at se dette, så læg først mærke til at

Bayes' Theorem Bevis ii

$E = E \cap S = E \cap (F \cup \overline{F}) = (E \cap F) \cup (E \cap \overline{F})$. Ydermere, $E \cap F$ og $E \cap \overline{F}$ er *disjoint* (usammenhængende), fordi hvis $x \in E \cap F$ og $x \in E \cap \overline{F}$, så $x \in F \cap \overline{F} = \emptyset$.

Konsekvent må $p(E) = p(E \cap F) + p(E \cap \overline{F})$. Vi har allerede vist at $p(E \cap F) = p(E|F)p(F)$. Derudover, har vi at $p(E|\overline{F}) = p(E \cap \overline{F})/p(\overline{F})$, som viser os at $p(E \cap \overline{F}) = p(E|\overline{F})p(\overline{F})$. Det følger nu at:

$$p(E) = p(E \cap F) + p(E \cap \overline{F}) = p(E|F)p(F) + p(E|\overline{F}p(\overline{F})$$

For at færdiggøre beviset plopper vi bare vores nye $p(E)$ ind og får:

$$p(F|E) = \frac{p(E|F)p(F)}{p(E|F)p(F) + p(E|\overline{F})p(\overline{F})}$$

Theorem 2

Theorem (Generalized Bayes' Theorem)

Suppose that E is an event from a sample space S and that F_1, $F_2, \ldots, F_n$ are mutually exclusive events such that $\bigcup_{i=1}^{n} F_i = S$. Assume that $p(E) \neq 0$ and $p(F_i) \neq 0$, for $i = 1, 2, \ldots, n$. Then

$$p(F_j|E) = \frac{p(E|F_j)p(F_j)}{\sum_{i=1}^{n} p(E|F_i)p(F_i)}$$

Reflektion

- Hvad er Bayes' Theorem?
- Hvad kan det bruges til?
- Hvornår?
- Hvorfor er $E = E \cap (F \cup \overline{F})$?

Expected Value

- **Expected Value** af en random variable er summen over alle elementer i et sample space af produktet af sandsynligheden af elementet og værdien af den tilfældelige variable på dette element. Konsekvent er expected value et vægtet gennemsnit af værdierne af en random variable.

Definition 1

Definition

The *expected value*, also called the *expectation* or *mean*, of the random variable X on the sample space S is equal to

$$E(X) = \sum_{s \in S} p(s)X(s)$$

The *deviation* of X at $s \in S$ is $X(s) - E(X)$, the difference between the value of X and the mean of X.

Læg mærke til at når sample space S har n elementer, $S = \{x_1, x_2, \ldots, x_n\}$, $E(X) = \sum_{i=1}^{n} p(x_i)X(x_i)$

Eksempel 1

- Nu er vi tilbage til nogle lidt nemmere opgaver.
- Så her er der også nogle skjulte svar.
- Lad X være tallet der kommer up når en terning bliver rullet. Hvad er expected value af X? (Næste slide indeholder sandsynligheden, ikke $E(X)$)

Eksempel 1

- Nu er vi tilbage til nogle lidt nemmere opgaver.
- Så her er der også nogle skjulte svar.
- Lad X være tallet der kommer up når en terning bliver rullet. Hvad er expected value af X? (Næste slide indeholder sandsynligheden, ikke $E(X)$)
- $p(X = r) = 1/6, r = 1, 2, \ldots 6$

Eksempel 1

- Nu er vi tilbage til nogle lidt nemmere opgaver.
- Så her er der også nogle skjulte svar.
- Lad X være tallet der kommer up når en terning bliver rullet. Hvad er expected value af X? (Næste slide indeholder sandsynligheden, ikke $E(X)$)
- $p(X = r) = 1/6, r = 1, 2, \ldots 6$
- $E[X] = \frac{1}{6} \cdot 1 + \frac{1}{6} \cdot 2 + \cdots + \frac{1}{6} \cdot 6 = \frac{21}{6} = \frac{7}{2}$

Theorem 1

Theorem

If X is a random variable and $p(X = r)$ is the probability that $X = r$, so that $p(X = r) = \sum_{s \in S, X(s) = r} p(s)$, then

$$E[X] = \sum_{r \in X(S)} p(X = r) r$$

- a.k.a., den langt nemmere måde at udregne $E[X]$ på.
- vent? er der en forskel? er oprigtigt i tvivl

Theorem 1 Bevis

Proof.

(Endelig et bevis så kort at vi kan have bevis-rammen tilbage!) Antag at X er en random variable med værdimængde $X(S)$, og let $p(X = r)$ være sandsynligheden at en random variable X tager vrædien r. Konsekvent, $p(X = r)$ er summen af sandsynlighederne s, således at $X(s) = r$. Det følger at

$$E[X] = \sum_{r \in X(S)} p(X = r) r$$

□

Forskel

- Ifølge ChatGPT er forskellene at:
  - Definition 1 summerer over hele sample spacet, og udregner indsatsen af hvert udfald individuelt.
  - Teoremet summerer over de distinkte værdier som et random variable kan have.

Theorem 2

Theorem

The expected number of successes when n mutually independent ***Bernoulli Trials*** *are performed, where p is the probability of success on each trial, is np.*

Theorem 2 Bevis

Lad X være det random variable lig med antallet af sucesser i n trials. Af Teorem 2 i 7.2 ser vi at $p(X = k) = \binom{n}{k} p^k q^{n-k}$. Dermed har vi:

Theorem 2 bevis part 2

$$
\begin{aligned}
E(X) &= \sum_{k=1}^{n} kp(X=k) \text{ by Theorem 1} \\
&= \sum_{k=1}^{n} k\binom{n}{k} p^k q^{n-k} \text{ by Theorem 2 in 7.2} \\
&= \sum_{k=1}^{n} n\binom{n-1}{k-1} p^k q^{n-k} \text{ by Exercise 21 in 6.4} \\
&= np\sum_{k=1}^{n} \binom{n-1}{k-1} p^{k-1} q^{n-k} \text{ factoring np from each term} \\
&= np\sum_{j=0}^{n-1} \binom{n-1}{j} p^{j} q^{n-1-j} \text{ shifting index of summation with} j = k -\\
&= np(p+q)^{n-1} \text{ by the binomial theorem} \\
&= np \text{ because p + q = 1}
\end{aligned}
$$

Theorem 3

Theorem

If $X_i, i = 1, 2, \ldots, n$ with n a positive integer, are random variables on S, and if a and b are real numbers, then

- $E(X_1 + X_2 + \cdots + X_n) = E(X_1) + E(X_2) + \cdots E(X_n)$
- $E(aX + b) = aE(X) + b$

Theorem 3 Beviser

Part 1:

$$
\begin{aligned}
E(X_1 + X_2) &= \sum_{s \in S} p(s)(X_1(s) + X_2(s)) \\
&= \sum_{s \in S} p(s) X_1(s) + \sum_{s \in S} p(s) X_2(s) \qquad (1) \\
&= E(X_1) + E(X_2)
\end{aligned}
$$

Part 2:

$$
\begin{aligned}
E(aX + b) &= \sum_{s \in S} p(s)(aX(s) + b) \\
&= a \sum_{s \in S} p(s) X(s) + b \sum_{s \in S} p(s) \qquad (2) \\
&= aE(X) + b \text{ because } \sum_{s \in S} p(s) = 1
\end{aligned}
$$

Eksempel 4

- Brug teorem 3 til at finde expected value af summen af tallene der fremkommer når et par af terninger bliver ullet.
- Hint: Lad X_1 og X_2 være random variables, hvor $X_1((i,j)) = i$, og $X_2((i,j)) = j$, altså, X_1 er værdien på første terning, og X_2 på anden terning. Hvad er værdierne af X_1 og X_2?

Eksempel 4

- Brug teorem 3 til at finde expected value af summen af tallene der fremkommer når et par af terninger bliver ullet.
- Hint: Lad X_1 og X_2 være random variables, hvor $X_1((i,j)) = i$, og $X_2((i,j)) = j$, altså, X_1 er værdien på første terning, og X_2 på anden terning. Hvad er værdierne af X_1 og X_2?
- Herfra kan det ses at $E[X_1] = E[X_2] = 7/2 = (1+2+3+4+5+6)/6$.

Eksempel 4

- Brug teorem 3 til at finde expected value af summen af tallene der fremkommer når et par af terninger bliver ullet.
- Hint: Lad X_1 og X_2 være random variables, hvor $X_1((i,j)) = i$, og $X_2((i,j)) = j$, altså, X_1 er værdien på første terning, og X_2 på anden terning. Hvad er værdierne af X_1 og X_2?
- Herfra kan det ses at $E[X_1] = E[X_2] = 7/2 = (1+2+3+4+5+6)/6$.
- Ved teorem 3: $E[X_1 + X_2] = E[X_1] + E[X_2] = 7/2 + 7/2 = 7$

Expected Value In The Hatcheck Problem

- Husker I Hatcheck Problemet? Lad os finde expected value.
- Hint: Lad X være random variable der er lig med antallet af folk som får den korrekte hat fra tjekkeren. Lad X_i være random variable hvor $X_i = 1$ hvis den i'e person får den korrekte hat, og $X_i = 0$ ellers (independent variables foreshadowing much?)
- Dermed må $X = X_1 + X_2 + \ldots + X_n$. Næste: udregn sandsynlighed for $P(X_i = 1)$

Expected Value In The Hatcheck Problem

- Husker I Hatcheck Problemet? Lad os finde expected value.
- Hint: Lad X være random variable der er lig med antallet af folk som får den korrekte hat fra tjekkeren. Lad X_i være random variable hvor $X_i = 1$ hvis den i'e person får den korrekte hat, og $X_i = 0$ ellers (independent variables foreshadowing much?)
- Dermed må $X = X_1 + X_2 + \ldots + X_n$. Næste: udregn sandsynlighed for $P(X_i = 1)$
- $P(X_i = 1) = 1/n$. Hvad er så $E[X_i]$?

Expected Value In The Hatcheck Problem

- Husker I Hatcheck Problemet? Lad os finde expected value.
- Hint: Lad X være random variable der er lig med antallet af folk som får den korrekte hat fra tjekkeren. Lad X_i være random variable hvor $X_i = 1$ hvis den i'e person får den korrekte hat, og $X_i = 0$ ellers (independent variables foreshadowing much?)
- Dermed må $X = X_1 + X_2 + \ldots + X_n$. Næste: udregn sandsynlighed for $P(X_i = 1)$
- $P(X_i = 1) = 1/n$. Hvad er så $E[X_i]$?
- $E[X_i] = 1 \cdot p(X_i = 1) + 0 \cdot p(X_i = 0) = 1/n$. Hvad er så $E[X]$?

Expected Value In The Hatcheck Problem

- Husker I Hatcheck Problemet? Lad os finde expected value.
- Hint: Lad X være random variable der er lig med antallet af folk som får den korrekte hat fra tjekkeren. Lad X_i være random variable hvor $X_i = 1$ hvis den i'e person får den korrekte hat, og $X_i = 0$ ellers (independent variables foreshadowing much?)
- Dermed må $X = X_1 + X_2 + \ldots + X_n$. Næste: udregn sandsynlighed for $P(X_i = 1)$
- $P(X_i = 1) = 1/n$. Hvad er så $E[X_i]$?
- $E[X_i] = 1 \cdot p(X_i = 1) + 0 \cdot p(X_i = 0) = 1/n$. Hvad er så $E[X]$?
- $E[X] = E[X_1] + E[X_2] + \cdots + E[X_n] = n \cdot 1/n = 1$

Eksempel 7: Expected Number of Inversions in a Permutation i

- The ordered pair (i, j) is called an *inversion* in a permutation of the first n positive integers if $i < j$, but j precedes i in the permutation. For instance, there are six inversions in the permutation $3, 5, 1, 4, 2$. These are:
- $(1, 3), (1, 5), (2, 3), (2, 4), (2, 5), (4, 5)$
- Let $I_{i,j}$ be the random variable on the set of all permutations of the first n positive integers with $I_{i,j} = 1$ if (i, j) is an inversion of the permutation and $I_{i,j} = 0$ otherwise. It follows

Eksempel 7: Expected Number of Inversions in a Permutation ii

that if X is the random variable equal to the number of inversions in the permutation, then

$$\sum_{1 \leq i < j \leq n} I_{i,j}$$

- Note that it is equally likely for i to precede j in a randomly chosen permutation as it is for j to precede i. (To see this, note that there are an equal number
- Note that it is equally likely for i to precede j in a randomly chosen permutation, as it is for j to precede i. Therefore:
- $E[I_{i,j}] = 1 \cdot p(I_{i,j} = 1) + 0 \cdot p(I_{i,j} = 0) = 1 \cdot 1/2 + 0 = 1/2$

Eksempel 7: Expected Number of Inversions in a Permutation iii

- Because there are $\binom{n}{2}$ pairs i and j with $1 \le i < j \le n$ and by the linearity of expectations, we have

$$E[X] = \sum_{1 \le i < j \le n} E[I_{i,j}] = \binom{n}{2} \cdot \frac{1}{2} = \frac{n(n-1)}{4}$$

Average-case Computational Complexity

- Du kan se det som at udregne expected value af en random variable.
- Lad sample space af et ekseperiment være sættet af mulige inputs, $a_j, j = 1, 2, \ldots, n$, og lad X være det random variable der giver a_j antallet af operationer brugt af algoritmen når a_j er brugt som input. Vi giver sandsynligheden $p(a_j)$ to hvert input værdi. Så er average-case complexity af algoritmen:
- $E[X] = \sum_{j=1}^{n} p(a_j)X(a_j)$

Eksempel 10

- Antag at sandsynligheden for at en mønt lander på plat er p. Mønten bliver kastet indtil den rammer plat. Hvor mange gange skal man kaste mønten før man får plat?
- Sandsynligheden for plat er $p(P) = p$. Videre: $p(KP) = (1-p)p, p(KKP) = (1-p)^2 p$ etc.
- Generelt, med n antal kroner, er sandsynligheden $(1-p)^{n-1}p$.
- $$E[X] = \sum_{j=1}^{\infty} j \cdot p(X=j) = \sum_{j=1}^{\infty} j(1-p)^{j-1}p = p\sum_{j=1}^{\infty} j(1-p)^{j-1} = p \cdot \frac{1}{p^2} =$$

Geometric Distrbution

Et random variable hvor antallet af kast før en mønt bliver plat, er et random variable med **geometric distribution**

Definition

A random variable X has a *geometric distribution with parameter* p if $p(X = k) = (1 - p)^{k-1}p$ for $k = 1, 2, 3, \ldots,$ where p is a real number with $0 \leq p \leq 1$.

Theorem (Theorem 4)

If the random variable X has the geometric distribution with parameter p, then $E[X] = 1/p$

Definition (Definition 3)

The random variables X and Y on a sample space S are *independent* if

$$p(X = r_1 \text{ and } Y = r_2) = p(X = r_1) \cdot p(Y = r_2)$$

or in words, if the probability that $X = r_1$ and $Y = r_2$ equals the product of the probabilities that $X = r_1$ and $Y = R_2$, for all real numbers r_1 and r_2.

Eksempel 11 i

- Er random variables X_1 og X_2 fra eksempel 4 uafhængige (independent)?
- Eksempel 4 er eksemplet med terningekast.
- Lad $S = \{1, 2, 3, 4, 5, 6\}$, og lad $i \in S$ og $j < inS$. Fordi der er 36 forskellige udfald når parret af terninger er rullet, og hvert er lige sandsynligt, har vi:

$$p(X_1 = i \text{ and } X_2 = j) = 1/36$$

.

Eksempel 11 ii

- Ydermere, $p(X_1 = i = 1/6$ og $p(X_2 = j) = 1/6$, fordi sandsynligheden i findes på den første terning, og sandsynligheden at j findes på den anden terning er begge $1/6$, følger det at:

$$p(X_1 = i) \text{ and } p(X_2 = j) = \frac{1}{36}$$

og

$$p(X_1 = i)p(X_2 = j) = \frac{1}{6} \cdot \frac{1}{6} = \frac{1}{36}$$

- Dermed er X_1 og X_2 uafhængige.

Eksempel 12 i

- Vis at random variablerne X_1 og $X = X_1 + X_2$, hvor X_1 og X_2 er som defineret i eksempel 4, ikke er uafhængige.
- Se først at $p(X_1 = 1 \text{ og } X = 12 = 0$, fordi X_1 betyder at antallet der vises på den første terning er 1, hvilket indebærer at summet ikke kan være 12.
- På den anden hånd er $p(X_1 = 1) = 1/6$ og $p(X = 12) = 1/36$. Dermed $p(X_1 = 1 \text{ og } X = 12) \neq p(X_1 = 1) \cdot p(X = 12)$. Dermed er de ikke uafhængige.

Theorem 5

Theorem (Theorem 5)

If X and Y are independent random variables on a sample space S, then $E[XY] = E[X]E[Y]$.

For at bevise dette bruger vi nøgleobservationen at $XY = r$ er den disjunkte forening af hændelserne $X = r_1$ og $Y = r_2$ hvor $r_1 \in X(S)$ og $r_2 \in Y(S)$ med $r = r_1 r_2$. Vi har

Theorem 5 Bevis part 2

$$
\begin{aligned}
E[XY] &= \sum_{r \in XY(S)} r \cdot p(XY = r) \text{ by Theorem 1} \\
&= \sum_{r_1 \in X(S), r_2 \in Y(S)} r_1 r_2 \cdot p(X = r_1 \text{ og } Y = r_2) \\
&= \sum_{r_1 \in X(S)} \sum_{r_2 \in Y(S)} r_1 r_2 \cdot p(X = r_1 \text{ og } Y = r_2) \\
&= \sum_{r_1 \in X(S)} \sum_{r_2 \in Y(S)} r_1 r_2 \cdot p(X = r_1 \cdot Y = r_2) \\
&= \sum_{r_1 \in X(S)} \left(r_1 \cdot p(X = r_1) \cdot \sum_{r_2 \in Y(S)} r_2 \cdot p(Y = r_2) \right)
\end{aligned}
$$

Theorem 5 Proof Part 3

$$
\begin{aligned}
&= \sum_{r_1 \in X(S)} r_1 \cdot p(X = r_1) \cdot E(Y) \\
&= E(Y) \left(\sum_{r_1 \in X(S)} r_1 \cdot p(X = r_1) \right) \\
&= E(Y) \cdot E(X)
\end{aligned}
$$

Variance

Variansen er en variabel der bestemmer hvor meget en random variable spreder sig. Ikke at jeg forstår hvad det betyder, men altså.

Definition

Let X be a random variable on a sample space S. The *variance* of X, denoted by $V(X)$, is

$$V(X) = \sum_{s \in S} (X(s) - E(X))^2 p(s)$$

That is, $V(X)$ is the weighted average of the square of the deviation of X. The *standard deviation* of X, denoted $\delta(X)$, is defined to be $\sqrt{V(X)}$.

Theorem 6

Theorem (Theorem 6)

If X is a random variable on a sample space S, then $V(X) = E(X^2) - E(X)^2$.

Theorem 6 Bevis

Læg mærke til at:

$$\begin{aligned} V(X) &= \sum_{s \in S} (X(S) - E(X))^2 p(s) \\ &= \sum_{s \in S} X(s)^2 p(s) - 2E(X) \sum_{s \in S} X(s) p(s) + E(X)^2 \sum_{s \in S} p(s) \\ &= E(X^2) - 2E(X)E(X) + E(X)^2 \\ &= E(X^2) - E(X)^2 \end{aligned}$$

Corollary 1

Corollary

If X is a random variable on a sample space S and $E(X) = \mu$, then $V(X) = E((X - \mu)^2)$.

Corollary 1 Proof

$$
\begin{aligned}
E((X-\mu)^2) &= E(X^2 - 2\mu X + \mu^2) \\
&= E(X^2) - E(2\mu X) + E(\mu^2) \\
&= E(X^2) - 2\mu E(X) + E(\mu^2) \\
&= E(X^2) - 2\mu E(X) + \mu^2 \text{ because } \mu^2 \text{ is a constant} \\
&= E(X^2) - 2\mu^2 + \mu^2 \\
&= E(X^2) - \mu^2 \\
&= V(X)
\end{aligned}
$$

Eksempel 15: Variance of the Value of a Die

- Hvad er variansen af random variable X, hvor X er tallet der kommer op når en terning bliver kastet?
- Svar: $V(X) = E(X^2) - E(X)^2$. Vi ved at $E(X) = 7/2$. $E(X^2) = \frac{1}{6}(1^2 + 2^2 + \cdots + 6^2) = \frac{91}{6}$.
- Dermed:
- $V(X) = \frac{91}{6} - \left(\frac{7}{2}\right)^2 = \frac{35}{12}$

Bienaymé's Formula

Theorem

If X and Y are two independent random variables on a sample space S, then $V(X+Y) = V(X) + V(Y)$. Furthermore, if $X_i, i = 1, 2, \ldots, n$, with n a positive integer, are pairwise independent random variables on S, then
$V(X_1 + X_2 + \cdots + X_n) = V(X_1) + V(X_2) + \cdots + V(X_n)$

Bienayme's Formula Bevis

Vi ved at $V(X+Y) = E((X+Y)^2) - E(X+Y)^2$ Dermed:

$$
\begin{aligned}
V(X+Y) &= E(X^2 + 2XY + Y^2 + Y^2) - (E(X) + E(Y))^2 \\
&= E(X^2) + 2E(XY) + E(Y^2) - E(X)^2 - 2E(X)E(Y) - E(Y)^2
\end{aligned}
$$

Fordi X og Y er uafhængige, ved vi at $E(XY) = E(X)E(Y)$, og dermed:

$$
\begin{aligned}
V(X+Y) &= (E(X^2) - E(X)^2) + E(Y^2) - E(Y)^2) \\
&= V(X) + V(Y)
\end{aligned}
$$

Chebyshev's Inequality

Theorem

Let X be a random variable on a sample space S with proabbility functino p. If r is a positive real number, then

$$p(|X(s) - E(X)| \geq r) \leq V(X)/r^2$$

Chebyshev's Inequality Bevis

Lad A være hændelsen:

$$A = \{s \in S \,||X(S) - E(X)| \geq r\} \tag{3}$$

Vi vil gerne bevise at $p(A) \leq V(X)/r^2$.

$$\begin{aligned}
V(X) &= \sum_{s \in S} (X(s) - E(X))^2 p(s) \\
&= \sum_{s \in A} (X(s) - E(X))^2 p(s) + \sum_{s \notin A} (X(s) - E(X))^2 p(s) \\
V(X) &\geq \sum_{s \in A} (X(s) - E(X))^2 p(s) \\
&= \sum_{s \in A} r^2 p(s) \\
&= r^2 p(A) \\
\frac{V(X)}{} &\geq p(A)
\end{aligned} \tag{4}$$

Chernoff Bounds

Chernoff Bounds

- Chernoff Bounds er bounds hvorpå du kan få et meget mere præcist bound end ved f.eks. Chebyshev's Inequality.

Let $X_1 + X_2 + \cdots + X_n$, where X_i takes the value 1 with the probability p_i, and the value 0 otherwise. $X = X_1 + X_2 + \cdots + X_n$. Assume that $\mu \geq E[X]$. Then, for any $\delta > 0$, we have

$$Pr[X > (1+\delta)\mu] < \left[\frac{e^{\delta}}{(1+\delta)^{(1+\delta)}}\right]^{\delta}. \tag{5}$$

Chernoff Bounds Bevis

Læg først mærke til at for enhver $t > 0, p(X > (1+\delta)\mu) = p(e^{tX} > e^{t(1+\delta)\mu})$, da funktionen $f(x) = e^{tx}$ er monoton i x.

Som det næste bruger vi nogle egenskaber for expectation. For et random variable, Y, har vi $\gamma p(Y > \gamma) \leq E[Y]$, af definition. Hvis vi sætter dette sammen, får vi:

$$p(X > (1+\delta)\mu) = p\left[e^{tX} > e^{t(1+\delta)\mu}\right] \leq e^{-t(1+\delta)\mu} E\left[e^{tX}\right] \tag{6}$$

Som det næste skal vi bounde $E[e^{tX}]$.
$E[e^{tX}] = E\left[e^{t\sum_i X_i}\right] = E[\Pi_i e^{tX_i}]$. For uafhængige variabler Y og Z, $E[YZ] = E[Y]E[Z]$. Dermed: $E[\pi_i e^{tX_i}] = \pi_i E[e^{tX_i}]$.

Chernoff Bevis part 2

e^{tX_i} er e^t med sandsynligheden p_i, og ellers e^0. Dermed kan vi bounde dens expectation med:

$$E\left[e^{tX_i}\right] = p_i e^t + (1 - p_i) = 1 + p_i(e^t - 1) \leq e^{p_i(e^t-1)}, \qquad (7)$$

hvor den sidste ulighed kommer fra at $1 + \alpha \leq e^{\alpha}$, så længe at $\alpha \geq 0$. Hvis vi kombinerer ulighederne får vi:

$$p(X > (1+\delta)\mu) \leq e^{-t(1+\delta)\mu} E\left[e^{tX}\right] = e^{-t(1+\delta)\mu} \Pi_i E\left[e^{tX_i}\right] \leq e^{-t(1+\delta)\mu} \Pi_i e^{p \qquad (8)$$

Givet at $t = \ln(1 + \delta)$ får vi bounde.

Theorem

Let $X, X_1, X_2, \ldots, X_n$ *and* μ *be defined as above. Then for any* $1 > \mu > 0$*, we have*

$$p(X < (1-\delta)\mu) < e^{-\frac{1}{2}\mu\delta^2} \tag{9}$$

Weekly Note 3

Indicator Random Variables

Lad S være sample space, lad A være en hændelse i S. Lad X_A være en random variable som er 1 hvis A hænder, og 0 ellers. Dermed:

$$X_A(s) = \begin{cases} 1 & \text{If } s \in A \\ 0 & \text{If } s \notin A \end{cases} \tag{10}$$

Vi kalder X_a indicator variable for A.

Theorem 1

Theorem

Let A be an event in a sample spce S and let X be the indicator variable for A. Suppose the probability that A occurs is $p(p(A) = p)$. Then expected value and variance of X is given by $E(X) = p, V(X) = p(1 - p)$

Theorem 1 Bevis i

Lad os starte med expected value:

$$\begin{aligned} E(X) &= 0 \cdot p(X=0) + 1 \cdot p(X=1) \\ &= 0 + 1 \cdot p \\ &= p \end{aligned} \tag{11}$$

Ved variansen bruger vi formlen $V(X) = E(X^2) - (E(X))^2$. Dette er nemt at gøre her, hvor hver indicator variable X $X(s) = X(s)^2$ for alle $s \in S$. Nu får vi:

Theorem 1 Bevis ii

$$
\begin{aligned}
V(X) &= E(X^2) + (E(X))^2 \\
&= E(X) - (E(X))^2 \\
&= E(X)(1 - E(X)) \\
&= p(1 - p)
\end{aligned}
\tag{12}
$$

Theorem 2

Theorem

Consider n independent Bernoulli trials each of which have probability of success equal to p. Let X be the random variable that denotes the number of successes in these n trials. Then $E(X) = np$ and $V(X) = np(1-p)$.

Theorem 2 Bevis

Vi har allerede bevist dette, dog ikke ved brug af indicator random variables.

Lad $X_i, i = 1, 2, \ldots, n$ være indicator random varibl for hændelsen at den i'e trial eer en success (så $p(X_i = 1) = p$ for all $i = 1, 2, \ldots, n)$. Så er $X = \sum_{i=1}^{X} X_i$. Ved linearity of expectation, får vi $E(X) = \sum_{i=1}^{n} E(X_i) = \sum_{i=1}^{n} p = np$. Ydermere, da X_i og X_j er uafhængige random variables, for alle distinkte i, j, har vi at $V(X) = \sum_{i=1}^{n} V(X_i) = \sum_{i=1}^{n} p(1 - p) = np(1 - p)$

Markov's Inequality

Theorem (Markov's Inequality)

For any non-negative random variable X on a sample space S,

$$p(X \geq t) \leq \frac{E(X)}{t} \tag{13}$$

Markov's Bevis

$E(X) = \sum_{i \in X(S)} p(X = i)i$. Siden $X(S) > 0$ for alle $s \in S$, får vi

$$E(X) \geq \sum_{i \in X(S), i \geq t} p(X = i)i \geq t \sum_{i \in X(S), i \geq t} p(X = i) = p(X \geq t)t. \tag{14}$$

First Moment Principle

Theorem

If $E(X) \leq t$, *then* $p(X \leq t) > 0$

First Moment Principle Bevis

Antag at $p(X \leq t) = 0$. Så har vi at:

$$E(X) = \sum_{i \in X(S)} p(X = i) i \geq \sum_{i \in X(S), i > t} p(X = i) i > t \sum_{i \in X(S), i \geq t} p(X = i) = t \tag{15}$$

, hvor vi har brugt at $p(X \leq t) = 0$