SOK-2009 – Høst 2023 - Arbeidskrav 1

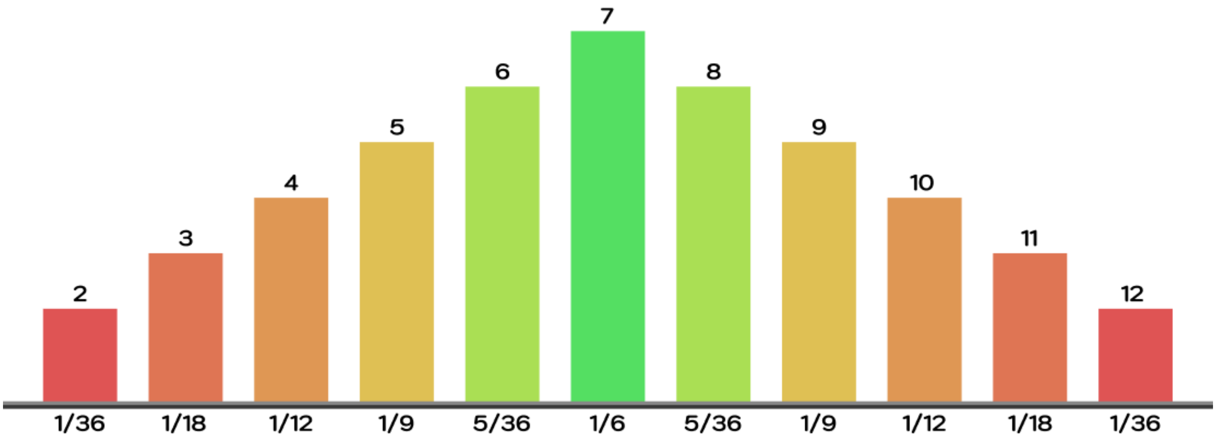
Kandidat nr. 12

Oppgave 6)

Du får i oppgave å lage et spill til kasino. Hver gang spillerne får et tall x på to terninger, i.e. $T_2 = x$ tjener spilleren en bonus (B) på 10 000 kroner.

1. En formel for forventet premie B om det bare er et utfall som gir premie er $E[B] = p(B)B$; kan du sette opp denne formelen dersom to eller N utfall gir premie.

I dette spillet kaster spillerne to terninger og får bonus (B) når summen av disse terningene er lik x. Vi vet at to terninger har 36 ulike kombinasjoner som kan gi summen 2 til 12. Nedenfor er sannsynligheten for disse kombinasjonene visualisert.



For å finne formelen for forventet premie ($E[B]$) kan vi bruke sannsynligheten for hvert av utfallene. Dermed blir formelen for forventet premie ($E[B]$) om det er to utfall som gir premier, gitt ved:

$$E[B] = p(B_1)B + p(B_2)B$$

Forventet premie (E[B]) er lik sannsynligheten for sum 1 ($p(B_1)$) ganget med bonusen (B) pluss sannsynligheten for sum 2 ($p(B_2)$) ganget med bonusen (B).

La oss si at når summen av to terninger er lik 5 eller 11 vil dette gi spilleren bonusen på 10 000 kroner. Da får vi:

$$1\,666.66 = \frac{1}{9} \times 10\,000 + \frac{1}{18} \times 10\,000$$

Her kan vi se at den forventede premien er lik 1 666.66 kroner når summen av to terninger er 5 eller 11.

Hvis det er N antall utfall som gir bonus (B), kan formelen ovenfor generaliseres på følgende måte:

$$E[B] = p(B_1)B + p(B_2)B + \ldots + p(B_n)B$$

2. Hva er minste spillet kan koste samtidig som kasino ikke taper penger (kasinoet har en kjempestor egenkapital og kan låne så mye de vil i banken).

For at kasinoet skal tjene penger må inntektene fra spillet være lik eller større enn utgiftene over tid. I dette kasinoet vil inntektene (I) være lik summen av antall spillere (S) ganget med innsatsen (C) per spiller: $I = S \times C$. Utgiftene (K) for dette kasinoet vil være lik summen av alle som vinner (V) ganget med bonusen (B): $K = V \times B$. For at kasinoet ikke skal tape penger må altså følgende formel være oppfylt: $I \geq K$.

Siden inntektene (I) er lik antall spillere (S) ganget med innsatsen (C), må innsatsen (C) være større eller lik den forventede premien per spiller.

La oss si at spillerne vinner når de kaster to seksere ($T_2 = 12$) og da får utbetalt bonusen (B = 10 000). Vi kan se fra modellen på forrige side at sannsynligheten for å vinne er dermed $\frac{1}{36}$.

Den forventede premien er gitt ved: $277.777778 = \frac{1}{36} \times 10\ 000$.

For å si det på en enklere måte vil i gjennomsnitt hver 36 spiller vinne bonusen på 10 000 kroner i dette spillet. Med 36 spillere og 1 vinner får vi følgende formel:

$I = 36 \times 277.77778 = 10\ 000.0001$ kr

$K = 1 \times 10\ 000 = 10\ 000.00$ kr

Da ser vi at formelen for at kasinoet ikke taper penger er oppfylt, altså:

$I \geq K\ (10\ 000.0001 \geq 10\ 000.00\ kr)$.

Det minste spillet kan koste samtidig som kasinoet ikke taper penger er $277.77 \approx 278$ kr

3. Kan du sette opp en B for å få et eller flere tall, en pris for deltakelse D for å spille slik at spillerne vinner med sannsynlighet: A. $\frac{1}{36}$ B. $\frac{2}{36}$ C. $\frac{1}{2}$

For å finne en B for å få et eller flere tall, og en pris for deltakelse D kan vi benytte figuren fra oppgave 1.

A – spilleren vinner med sannsynlighet $\frac{1}{36}$

La oss sette vinnertallet til summen 12, noe som har en sannsynlighet på $\frac{1}{36}$, da får vi at:

$$D = \frac{1}{36}B$$

Så kan vi sette inn en premie på 10 000 kr, og da får vi at prisen for deltakelse må være ca. 277,77 kr for at kasinoet ikke skal tape penger over tid.

$$277{,}77 \approx \frac{1}{36} \times 10\ 000$$

B – spilleren vinner med sannsynlighet $\frac{2}{36}$

La oss sette vinnertallet til summen 12 og 2 som har sannsynligheter på $\frac{1}{36} + \frac{1}{36} = \frac{2}{36}$, da får vi at:

$$D = \frac{2}{36}B$$

Så kan vi sette inn en premie på 10 000 kr, og da får vi at prisen for deltakelse må være ca. 555,55 kr for at kasinoet ikke skal tape penger over tid.

$$555{,}55 \approx \frac{2}{36} \times 10\ 000$$

C – spilleren vinner med sannsynlighet $\frac{1}{2}$.

La oss sette vinnertallet til summen 3, 5, 7, 9 og 11 som har sannsynligheter på $\frac{1}{18} + \frac{1}{9} + \frac{1}{6} + \frac{1}{9} + \frac{1}{18} = \frac{1}{18} + \frac{2}{18} + \frac{3}{18} + \frac{2}{18} + \frac{1}{18} = \frac{9}{18} = \frac{1}{2}$, da får vi at:

$$D = \frac{1}{2}B$$

Så kan vi sette inne en premie på 10 000 kr, og da får vi at prisen for deltakelse må være 5 000 kr for at kasinoet ikke skal tape penger over tid.

$$5\ 000 = \frac{1}{2} \times 10\ 000$$

4.1. En gruppe kunder har tapt mange penger og er lei seg. De klager og sier det er urettferdig at kasinoet ditt tar mer betalt for å spille en forventet utbetaling. Hvordan forklarer du dette til kundene?

Kasinoer er utformet for å tjene penger, og dette gjør de ved å tilby spill med en forventet utbetaling. Det vil si at alle spillene kasinoet tilbyr har en innebygd fordel for kasinoet, som over tid gjør at kasinoet tjener penger. Denne fordelen er nødvendig for at kasinoet skal tjene penger og visst kasinoet ikke hadde hatt en slik fordel, ville de ikke kunne betale for driftsutgiftene, inkludert premier til kundene sine over lengere tid.

4.2. Eieren av kasinoet er lei seg fordi banken tar høyere rente av lånet hans en det banken betaler i innskuddsrente. Hvordan forklarer du til eieren at banken må ta rente.

Når banker låner ut penger tar de en risiko, de er aldri 100 prosent sikker på at de får tilbake pengene sine. En høyere rente på lånet vil hjelpe til med å kompensere for denne risikoen. Banker har også ansatte og andre administrative kostnader, og disse utgiftene dekkes i stor grad av rentene på lån, og som enhver annen bedrift er bankene interessert i å tjene penger. Inntektene fra rentene på lån utgjør en viktig del av deres fortjeneste.

En gruppe på 5 studenter $i \in \{1,2,3,4,5\}$ går på kasino. Vi skriver poengsummen i statistikkfaget deres på skalaen $p_i \in [1,100]$ og gevinsten på kasino $s_i \in [-100,100]$. Data er som følger, $\{p_i, s_i\}_i$: $\{10,-50\}_1, \{78,-100\}_2, \{98,100\}_3, \{60,50\}_4, \{30,25\}_5$.

5. Kan du sette opp gjennomsnitt og median for p_i og s_i. Er det noe type tall i utvalget? Sett også opp varians, standardavvik for s_i, p_i og korrelasjon og kovarians mellom s_i, p_i.

Gjennomsnitt:

Gjennomsnitt er en middelverdi av alle dataene. Vi finner gjennomsnittet ved å summere alle dataene og dividere på det totale antall data (Mattematikk.org, I.D.).

$$p_i = \frac{10 + 78 + 98 + 60 + 30}{5} = 55.2$$

$$s_i = \frac{(-50) + (-100) + 100 + 50 + 25}{5} = 5$$

Median:

Medianen finner vi ved først å sortere alle dataene i stigende rekkefølge, og finne det midterste tallet (Mattematikk.org, I.D.).

$p_i = \{10,30,60,78,98\} = 60$

$s_i = \{-100,-50,25,50,100\} = 25$

Typetall

Typetallet er den mest «typiske» observasjonen i en datasamling, og det er det tallet som forekommer flest ganger (Mattematikk.org, I.D.).

Ingen av de samme tallene forekommer flere ganger blant observasjonene, vi kan derfor ikke si at det foreligger et typetall.

Varians

Varians er definert som summen av kvadratet av hver observasjons avstand fra gjennomsnittet dividert med det totale antallet observasjoner. Variansen finner vi ved å (1) regne ut gjennomsnittet, (2) regne ut forskjellene mellom gjennomsnittet og hver av tallene, (3) kvadrere forskjellene, (4) summere kvadratene av forskjellene, (5) dividere summen med det totale antallet observasjoner (Mattematikk.org, I.D.).

(1) $p_i = 55.2$ og $s_i = 5$.

(2) $p_i = 10 - 55.2 = -45.2 \mid 78 - 55.2 = 22.8 \mid 98 - 55.2 = 43.8 \mid$

$60 - 55.2 = 4.8 \mid 30 - 55.2 = -25.2$

$s_i = (-50) - 5 = -55 \mid (-100) - 5 = -105 \mid 100 - 5 = 95 \mid$

$50 - 5 = 45 \mid 25 - 5 = 20$

(3) $p_i = (-45.2)^2 = 2043.04 \mid 22.8^2 = 519.84 \mid 43.8^2 = 1918.44 \mid 4.8^2 = 23.04 \mid$

$(-25.2)^2 = 635.04$

$s_i = (-55)^2 = 3025 \mid (-105)^2 = 11025 \mid 95^2 = 9025 \mid 45^2 = 2025 \mid$

$20^2 = 400$

(4) $p_i = 2043.04 + 519.84 + 1918.44 + 23.04 + 635.04 = 5139.4$

$s_i = 3025 + 11025 + 9025 + 2025 + 400 = 25500$

(5) $p_i = \frac{5139.4}{5} = 1027.88$

$s_i = \frac{25500}{5} = 5100$

Standardavvik

Variansen er ikke så lett å tolke, så etter å ha regnet den ut tar vi kvadratroten av variansen, og tallet vi får da kalles for standardavviket. Dette er et forventet avvik fra gjennomsnittet (Mattematikk.org, I.D.).

$p_i = \sqrt{1027.88} \approx 32.06$

$s_i = \sqrt{5100} \approx 71.41$

Korrelasjon

Korrelasjon er et matematisk mål på sammenhengen mellom to variable størrelser. Korrelasjonstallet ρ er et tall som ligger mellom -1 og 1, hvor fortegnet angir retningen til sammenhengen. Det er absoluttverdien til ρ som sier noe om hvor sterk sammenhengen er. Jo større ρ er jo sterkere er sammenhengen, og jo bedre linearitet er det mellom de to datasettene (Kjøbli, 2018).

For å beregne korrelasjonen kan vi bruke følgende formel:

$$\rho = \frac{\sum(s_i - \bar{s})(p_i - \bar{p})}{\sqrt{\sum(s_i - \bar{s})^2}\sqrt{\sum(p_i - \bar{p})^2}}$$

Her er ρ korrelasjonskoeffisienten, $\bar{s}\ og\ \bar{p}$ er gjennomsnittene for $s_i\ og\ p_i\ og\ \sum$, betyr summen av alle observasjonene. Vi har allerede funnet flere av disse verdiene i oppgavene tidligere og da får vi:

$$0.3463 \approx \frac{(-55)(-45.2) + (-105)(22.8) + (95)(43.8) + (45)(4.8) + (20)(-25.2)}{\sqrt{25\,500}\sqrt{5\,139.4}}$$

En korrelasjon tilsvarende 0.3463 tilsier en lav positiv korrelasjon mellom poeng i statistikkfaget (p_i) og gevinst på kasino (s_i).

Kovarians

Kovarians er innen statistikk et mål på sammenhengen mellom to ulike størrelser eller målbare fenomen (Frøslie, 2020).

For å bergene kovarians kan vi bruke følgende formel:

$$\mathrm{Cov}(\mathrm{p_i}, \mathrm{s_i}) = \frac{1}{n}\sum_{i=1}^{n}(p_i - \bar{p})(s_i - \bar{s})$$

Her er n antall observasjoner, p_i poengsum i statistikkfaget, s_i gevinsten på kasinoet, $\bar{p}$ gjennomsnittet av poengsummen i statistikkfaget og $\bar{s}$ gjennomsnittet av gevinsten på kasinoet. Vi har alle verdiene vi trenger fra tidligere oppgaver og da får vi:

$$Cov(p_i, s_i) = \frac{1}{5}((-55)(-45.2) + (-105)(22.8) + (95)(43.8) + (45)(4.8) + (20)(-25.2))$$

$$Cov(p_i, s_i) = 793$$

En kovarians tilsvarende 793 tilsier en positiv samvariasjon mellom poeng i statistikkfaget (p_i) og gevinst på kasino (s_i).

Referanseliste

Mattematikk.org. (I.D.). *Gjennomsnitt, median og typetall.* Hentet 13. oktober 2023 fra https://www.matematikk.org/artikkel.html?tid=154336&within_tid=154329

Mattematikk.org. (I.D.). *Varians og standardavvik.* Hentet 13. oktober 2023 fra https://www.matematikk.org/artikkel.html?tid=154338&within_tid=154329

Kjøbli, E. (12. oktober, 2018). *Korrelasjon.* NTNU. Hentet 19.09.2023 fra https://www.ntnu.no/wiki/display/medtekipedia/Korrelasjon

Frøslie, K.F. (30. januar, 2020). *Kovarians.* SNL. Hentet 19.09.2023 fra https://snl.no/kovarians