MVE090 LP4 2021 Föreläsning 3-4

Kapitel 3. “Discrete Distributions”

Definition 3.1. En <u>stokastisk variabel</u> X är en funktion $X : S \to \mathbb{R}$ från ett utfallsrum S till de reella talen $\mathbb{R}$.

Till varje utfall $\omega \in S$ ordnar alltså den stokastiska variabeln X ett reellt tal $X(\omega)$.

Exempel 2.1. (Forts.) För kast med två tärningar är summan av vad tärningarna visar ett exempel på en stokastisk variabel, dvs.

$$S = \big\{\{(i,j) : i,j \in \{1,2,3,4,5,6\}\big\} \quad \text{och} \quad X((i,j)) = i+j.$$

Definition 3.2. En stokastisk variabel X är <u>diskret</u> om den har antingen ändligt många eller uppräkneligt oänligt många (olika) möjliga värden.

Exempel på ändliga mängder är $\{1,2,\ldots,n\}$ etc. Exempel på uppräkneligt oändliga mängder är heltal $\mathbb{N} = \{0,1,2,\ldots\}$ eller $\mathbb{Z} = \{0,\pm 1,\pm 2,\ldots\}$ och rationella tal $\mathbb{Q}$.

Definition 3.3. För en diskret stokastisk variabel X definieras <u>frekvensfunktionen</u> eller <u>täthetsfunktionen</u> $f(x) = f_X(x)$ [två olika benämningar för samma funktion som på engelska kallas “(probability) density/mass function”] genom

$$f(x) = f_X(x) = \mathbf{P}(X = x) = \mathbf{P}(\{\omega \in S : X(\omega) = x\}).$$

Exempel 2.1. (Forts.) För X summan av två balanserade tärningar är

$$f_X(x) = \begin{cases} 1/36 & \text{för} \quad x=2 \ \text{och} \ x=12 \\ 2/36 & \text{för} \quad x=3 \ \text{och} \ x=11 \\ 3/36 & \text{för} \quad x=4 \ \text{och} \ x=10 \\ 4/36 & \text{för} \quad x=5 \ \text{och} \ x=9 \\ 5/36 & \text{för} \quad x=6 \ \text{och} \ x=8 \\ 6/36 & \text{för} \quad x=7 \end{cases}.$$

Här använder vi klassiska sannolikhetsformeln enligt (1) i kapitel 2. Tex. innehåller händelsen $A = \{X = 4\}$ = “summan av tärningarna = 4” de tre utfallen $(1,3)$, $(2,2)$ och $(3,1)$ med sannolikhet $1/36$ vardera så att $\mathbf{P}(X = 4) = 3/36$.

Definition 3.4. För en stokastisk variabel X definieras <u>fördelningsfunktionen</u> $F(x) = F_X(x)$ (engelska “cummulative distribution function”) genom

$$F(x) = F_X(x) = \mathbf{P}(X \le x) = \mathbf{P}(\{\omega \in S : X(\omega) \le x\}).$$

Exempel 2.1. (Forts.) För X summan av två balanserade tärningar är

$$F_X(x) = \begin{cases} 0 & \text{för} \quad x < 2 \\ 1/36 & \text{för} \quad 2 \le x < 3 \\ 3/36 & \text{för} \quad 3 \le x < 4 \\ 6/36 & \text{för} \quad 4 \le x < 5 \\ 10/36 & \text{för} \quad 5 \le x < 6 \\ 15/36 & \text{för} \quad 6 \le x < 7 \\ 21/36 & \text{för} \quad 7 \le x < 8 \\ 26/36 & \text{för} \quad 8 \le x < 9 \\ 30/36 & \text{för} \quad 9 \le x < 10 \\ 33/36 & \text{för} \quad 10 \le x < 11 \\ 35/36 & \text{för} \quad 11 \le x < 12 \\ 1 & \text{för} \quad 12 \le x \end{cases}.$$

Här använder vi vad vi vet om $f_X(x)$ från ovan. Tex. är

$$\begin{aligned} F_X(3.5) = \mathbf{P}(X \le 3.5) &= \mathbf{P}(\{X = 2\} \cup \{X = 3\}) \\ &= \mathbf{P}(X = 2) + \mathbf{P}(X = 3) = f_X(2) + f_X(3) = 3/36. \end{aligned}$$

Sats 3.5. För en diskret stokastisk variabel X med frekvensfunktion $f_X(x)$ är

1. $f_X(x) \ge 0$ för all x,
2. $\sum_{\text{alla } x} f_X(x) = 1$,
3. $\mathbf{P}(X \in C) = \sum_{\text{alla } x \in C} f_X(x)$ för $C \subseteq \mathbb{R}$.

Bevis. 1. Följer av att $f_X(x)$ är en sannolikhet.

2. Följer av att ta $C = \mathbb{R}$ i 3.

3. Analogt med hur jag visade $F_X(3.5) = 3/36$ för X summan av två tärningar ovan är

$$\mathbf{P}(X \in C) = \mathbf{P}\Big(\bigcup_{x \in C} \{X = x\}\Big) = \sum_{x \in C} \mathbf{P}(X = x).$$

Här följer andra likheten av 3 i Sats 2.2 eftersom $\{X = y\} \cap \{X = z\} = \emptyset$ för $y \ne z$: den stokastiska variabeln X kan inte vara lika med både y och z på en gång. □

Definition 3.6. <u>Väntevärdet</u> $\mathbf{E}(X)$ av en diskret stokastisk variabel X definieras

$$\mathbf{E}(X) = \sum_{\text{alla } x} x\, f_X(x).$$

Väntevärde (engelska “expectation) är tyngdpunkten för den diskreta grafen av frekvensfunktionen $f_X(x)$. Väntevärdet $\mathbf{E}(X)$ betecknas ofta μ eller (tydligare) μ_X.

Exempel 2.1. (FORTS.) För X summan av två balanserade tärningar inses lätt [studera uttrycket för $f_X(x)$ ovan] att tyngdpunkten för $f_X(x)$ är $\mathbf{E}(X) = 7$. Detta kan även räknas ut (utan behöva "inse" något) mha. definition 3.6:

$$\mathbf{E}(X)$$
$$= 2\cdot\tfrac{1}{36} + 3\cdot\tfrac{2}{36} + 4\cdot\tfrac{3}{36} + 5\cdot\tfrac{4}{36} + 6\cdot\tfrac{5}{36} + 7\cdot\tfrac{6}{36} + 8\cdot\tfrac{5}{36} + 9\cdot\tfrac{4}{36} + 10\cdot\tfrac{3}{36} + 11\cdot\tfrac{2}{36} + 12\cdot\tfrac{1}{36} = 7.$$

Sats 3.7. Väntevärdet av funktionen $h(X)$ av en diskret stokastisk variabel X är

$$\mathbf{E}(h(X)) = \sum_{\text{alla } k} h(k)\, f_X(k).$$

Exempel 2.1. (FORTS.) För X summan av två balanserade tärningar och

$$h(x) = \begin{cases} 0 & \text{om} \quad x \in \{2,4,6,8,10,12\} \quad \text{är jämn} \\ 1 & \text{om} \quad x \in \{3,5,7,9,11\} \quad \text{är udda} \end{cases}$$

är $\mathbf{E}(h(X)) = 1/2$ ty sannolikheten att X är jämn är $f_X(2) + f_X(4) + f_X(6) + f_X(8) + f_X(10) + f_X(12) = \ldots = 1/2$ så att sannolikheten att X är udda också är $1/2$. Detta kan även räknas ut (utan behöva inse något) mha. sats 3.7:

$$\mathbf{E}(h(X))$$
$$= 0\cdot\tfrac{1}{36} + 1\cdot\tfrac{2}{36} + 0\cdot\tfrac{3}{36} + 1\cdot\tfrac{4}{36} + 0\cdot\tfrac{5}{36} + 1\cdot\tfrac{6}{36} + 0\cdot\tfrac{5}{36} + 1\cdot\tfrac{4}{36} + 0\cdot\tfrac{3}{36} + 1\cdot\tfrac{2}{36} + 0\cdot\tfrac{1}{36} = \frac{1}{2}.$$

Sats 3.8. För två diskreta stokastiska variabler X och Y är

1. $\mathbf{E}(X) = c$ om X är en konstant stokastisk variabel $X = c$ för ett $c \in \mathbb{R}$,
2. $\mathbf{E}(c\,X) = c\,\mathbf{E}(X)$ för en konstant $c \in \mathbb{R}$,
3. $\mathbf{E}(X + Y) = \mathbf{E}(X) + \mathbf{E}(Y)$.

Bevis. 1. Enligt definition 3.6 är

$$\mathbf{E}(X) = \sum_{\text{alla } x} x\, f_X(x) = c\, f_X(c) = c \cdot 1 = c.$$

2. Genom ta $h(x) = c\,x$ i sats 3.7 och därefter använda definition 3.6 ser jag att

$$\mathbf{E}(c\,X) = \sum_{\text{alla } x} h(x)\, f_X(x) = \sum_{\text{alla } x} c\,x\, f_X(x) = c\,\mathbf{E}(X).$$

3. Detta kan ej bevisas förrän en bit in i kapitel 5 i M&A. □

Definition 3.9. Variansen $\mathbf{Var}(X)$ av en stokastisk variabel X definieras

$$\mathbf{Var}(X) = \mathbf{E}\big([X - \mathbf{E}(X)]^2\big) = \mathbf{E}((X - \mu_X)^2).$$

Standardavvikelsen av X definieras som $\sqrt{\mathbf{Var}(X)}$.

Variansen $\mathbf{Var}(X)$ betecknas ofta σ^2 eller σ_X^2 och standardavvikelsen σ eller σ_X.

$\mathbf{Var}(X)$ är ett kvadratiskt mått på hur utspridd grafen av frekvensfuntionen är kring tyngpunkten $\mathbf{E}(X)$ – stor varians betyder bred graf och liten varians smal graf koncentrerad kring $\mathbf{E}(X)$. σ_X betyder samma i verklig (ej kvadrerad) skala.

Sats 3.10. För två diskreta stokastiska variabler X och Y är

1. $\mathbf{Var}(X) = \sum_{\text{alla } x} (x - \mu_X)^2$,
2. $\mathbf{Var}(X) = \mathbf{E}(X^2) - (\mathbf{E}(X))^2 = \mathbf{E}(X^2) - \mu_X^2$,
3. $\mathbf{Var}(X) = 0$ om X är en konstant stokastisk variabel $X = c$ för ett $c \in \mathbb{R}$,
4. $\mathbf{Var}(c\,X) = c^2\,\mathbf{Var}(X)$ för en konstant $c \in \mathbb{R}$,
5. $\mathbf{Var}(X+Y) = \mathbf{Var}(X) + \mathbf{Var}(Y)$ om X och Y är oberoende.

Att X och Y är oberoende betyder att $\mathbf{P}(\{X \in C\} \cap \{Y \in D\}) = \mathbf{P}(X \in C)\,\mathbf{P}(Y \in D)$ för $C, D \subseteq \mathbb{R}$. Detta läres ut i kapitel 5 i M&A så fundera inte mer på det nu.

Föresten skrives ofta $\mathbf{P}(A, B)$ istf. $\mathbf{P}(A \cap B)$ i fortsättningen – de är synonyma.

Bevis. 1. Följer direkt av använda sats 3.7 med $h(x) = (x - \mu_X)^2$.

2. Genom använda alla tre delarna av sats 3.8 följer att

$$\begin{aligned}\mathbf{Var}(X) = \mathbf{E}((X-\mu_X)^2) &= \mathbf{E}\big(X^2 - 2\,\mu_X\,X + \mu_X^2\big) \\ &= \mathbf{E}(X^2) - 2\,\mu_X\,\mathbf{E}(X) + \mathbf{E}(\mu_X^2) = \mathbf{E}(X^2) - \mu_X^2.\end{aligned}$$

3. Enligt 1 i sats 3.8 är $\mathbf{E}(X) = c$ så att $(X - \mu_X)^2 = 0$ som har väntevärde $\mathbf{E}(0) = 0$.

4. Enligt 2 i sats 3.8 är $(c\,X - \mu_{cX})^2 = (c\,X - c\,\mu_X)^2 = c^2\,(X - \mu_X)^2$ som (enligt 2 i sats 3.8) har väntevärde $c^2\,\mathbf{E}((X-\mu_X)^2) = c^2\,\mathbf{Var}(X)$.

5. Detta kan ej bevisas förrän en bit in i kapitel 5 i M&A. $\square$

Exempel 2.1. (Forts.) För X summan av två balanserade tärningar är $\mathbf{E}(X^2)$

$2^2\cdot\frac{1}{36}+3^2\cdot\frac{2}{36}+4^2\cdot\frac{3}{36}+5^2\cdot\frac{4}{36}+6^2\cdot\frac{5}{36}+7^2\cdot\frac{6}{36}+8^2\cdot\frac{5}{36}+9^2\cdot\frac{4}{36}+10^2\cdot\frac{3}{36}+11^2\cdot\frac{2}{36}+12^2\cdot\frac{1}{36}$

som blir $\ldots = 329/6$ så att enligt sats

$$\mathbf{Var}(X) = \mathbf{E}(X^2) - (\mathbf{E}(X))^2 = \tfrac{329}{6} - 7^2 = \ldots = \tfrac{35}{6}.$$

Definition 3.11. Den <u>momentgenererande funktionen (MGF)</u> $m_X : [0, \infty) \to \mathbb{R}$ för en stokastisk variabel X definieras

$$m_X(t) = \mathbf{E}(\mathrm{e}^{tX}).$$

Sats 3.12. För en stokastisk variabel X gäller att

$$m_X^{(k)}(0) = \tfrac{d^k}{dt^k}\, m_X(t)\,\Big|_{t=0} = \mathbf{E}(X^k) \quad \text{för } \; k \in \mathbb{N}.$$

Bevis. $m_X^{(k)}(0) = \tfrac{d^k}{dt^k}\,\mathbf{E}(\mathrm{e}^{tX})\,\Big|_{t=0} = \mathbf{E}\big(\tfrac{d^k}{dt^k}\,\mathrm{e}^{tX}\big)\,\Big|_{t=0} = \mathbf{E}\big(X^k\,\mathrm{e}^{tX}\big)\,\Big|_{t=0} = \mathbf{E}(X^k).$ $\square$

Talet $\mathbf{E}(X^k)$ kallas moment av ordning k för den stokastiska variabeln X och $k \in \mathbb{N}$.

Definition 3.13. En stokastisk variabel X med möjliga värden $1, 2, 3\ldots$ och frekvensfunktion $f_X(k) = (1-p)^{k-1}p$ för $k = 1, 2, 3\ldots$ är geometriskt fördelad.

Geometrisk fördelning kallas även väntetidsfördelning ty om ett slumpmässigt experiment med sannolikhet p att lyckas utföres upprepat tills det lyckas första gången och X är antalet försök som behövs för det blir X geometriskt fördelad eftersom $X = k$ betyder att först misslyckas $k-1$ försök varefter experiment nummer k lyckas.

Exempel 3.1. En geometrisk fördelad stokastisk variabel X har enligt sats 3.7 och formeln för geometrisk summa momentgenererande funktion

$$\mathbf{E}(\mathrm{e}^{tX}) = \sum_{\text{alla } x} \mathrm{e}^{tx} f_X(x) = \sum_{k=1}^{\infty} \mathrm{e}^{tk}(1-p)^{k-1}p = \mathrm{e}^t p \sum_{k=1}^{\infty} [\mathrm{e}^t(1-p)]^{k-1} = \frac{\mathrm{e}^t p}{1-\mathrm{e}^t(1-p)}.$$

Mha. Mathematica beräknas därför lätt $m'_X(0) = \mathbf{E}(X)$

```
In[1]:= Simplify[D[Exp[t]*p/(1-Exp[t]*(1-p)),t]/.{t->0}]
```

Out[1]= $\frac{1}{p}$

och $m''_X(0) = \mathbf{E}(X^2)$

```
In[2]:= Simplify[D[D[Exp[t]*p/(1-Exp[t]*(1-p)),t],t]/.{t->0}]
```

Out[2]= $\frac{2-p}{p^2}$

så att enligt 2 i sats 3.10

$$\mathbf{Var}(X) = \mathbf{E}(X^2) - (\mathbf{E}(X))^2 = \tfrac{2-p}{p^2} - (\tfrac{1}{p})^2 = \tfrac{1-p}{p^2}.$$

Definition 3.14. En stokastisk variabel X med möjliga värden $0, 1, \ldots, n$ och frekvensfunktion $f_X(k) = \binom{n}{k}\, p^k\,(1-p)^{n-k}$ för $k = 0, 1, \ldots, n$ är binomialfördelad.

Om ett slumpmässigt experiment med sannolikhet p att lyckas utföres n gånger och X är antalet lyckade försök blir X binomialfördelad. Ty varje specifik sekvens av k lyckade och $n-k$ misslyckade försök har sannolikhet $p^k\,(1-p)^{n-k}$ och enligt sats 1.4 finns $\binom{n}{k}$ olika sådana sekvenser.

Exempel 3.2. För en binomialfördelad stokastisk variabel X är enligt sats 3.7

$$\mathbf{E}(\mathrm{e}^{tX}) = \sum_{\text{alla } x} \mathrm{e}^{tx} f_X(x) = \sum_{k=0}^{n} \mathrm{e}^{tk} \binom{n}{k} p^k (1-p)^{n-k} =_{\text{binomialteoremet}} \left(\mathrm{e}^t p + (1-p)\right)^n.$$

Mha. Mathematica beräknas därför lätt $m'_X(0) = \mathbf{E}(X)$

```
In[3]:= Simplify[D[(Exp[t]*p+(1-p))^n,t]/.{t->0}]
```

`Out[3]=` $n\,p$

och $m''_X(0) = \mathbf{E}(X^2)$

```
In[4]:= Simplify[D[D[(Exp[t]*p+(1-p))^n,t],t]/.{t->0}]
```

`Out[4]=` $n\,p\,(1+(n-1)\,p)$

så att enligt 2 i sats 3.10

$$\mathbf{Var}(X) = \mathbf{E}(X^2) - (\mathbf{E}(X))^2 = n\,p\,(1+(n-1)\,p) - (n\,p)^2 = n\,p\,(1-p).$$

Definition 3.15. En stokastisk variabel X med möjliga värden $\mathbb{N} = 0, 1, 2 \ldots$ och frekvensfunktion $f_X(k) = (\lambda^k/(k!))\,\mathrm{e}^{-\lambda}$ för $k \in \mathbb{N}$ är Poisson fördelad.

Poisson fördelningen förekommer i naturen: det slumpmässiga antal sönderfall ett stycke radioaktivt material har per tidsenhet har denna födelning.

Exempel 3.3. För en Poisson fördelad stokastisk variabel X är

$$m_X(t) = \sum_{k=0}^{\infty} \mathrm{e}^{tk}\,(\lambda^k/(k!))\,\mathrm{e}^{-\lambda} = \mathrm{e}^{-\lambda} \sum_{k=0}^{\infty} (\mathrm{e}^t\lambda)^k/(k!) = \mathrm{e}^{-\lambda}\mathrm{e}^{\mathrm{e}^t\lambda} = \mathrm{e}^{(\mathrm{e}^t-1)\lambda}$$

enligt Taylor utveckling, så att $m'_X(0) = \mathbf{E}(X)$ är

```
In[5]:= Simplify[D[Exp[(Exp[t]-1)*lambda],t]/.{t->0}]
```

`Out[5]=` *lambda*

och $m''_X(0) = \mathbf{E}(X^2)$ är

```
In[6]:= Simplify[D[D[Exp[(Exp[t]-1)*lambda],t],t]/.{t->0}]
```

`Out[6]=` *lambda* $(1 + $ *lambda*$)$

vilket ger

$$\mathbf{Var}(X) = \mathbf{E}(X^2) - (\mathbf{E}(X))^2 = \mathit{lambda}\,(1+\mathit{lambda}) - \mathit{lambda}^2 = \mathit{lambda}.$$

Definition 3.16. En stokastisk variabel X med möjliga värden $r, r+1, r+2, \ldots$ kallas negativt binomialfördelad om den har frekvensfunktion

$$f_X(k) = \binom{k-1}{r-1} p^{r-1}(1-p)^{(k-1)-(r-1)}p = \binom{k-1}{r-1} p^r (1-p)^{k-r} \quad \text{för } k = r, r{+}1, r{+}2, \ldots .$$

Om ett experiment med sannolikhet p att lyckas varje gång utföres till det lyckats r gånger och X är antalet experiment som behövs blir X negativt binomialfördelad. Ty $X = k$ betyder att först är en $\mathrm{Bin}(k-1,p)$ stokastisk variabel lika $r-1$ med sannolikhet $\binom{k-1}{r-1}\, p^{r-1}\,(1-p)^{(k-1)-(r-1)}$ och sedan lyckas experiment nummer k med sannolikhet p.

Definition 3.17. En stokastisk variabel X är <u>hypergeometriskt fördelad</u> om den har möjliga värden $\{\max[0,n-(N-r))],\ldots,\min[n,r]\}$ och frekvensfunktion

$$f_X(k) = \binom{r}{k}\binom{N-r}{n-k}\,/\,\binom{N}{n} \quad \text{för } k \in \{\max[0,n-(N-r))],\ldots,\min[n,r]\}.$$

Om man slumpmässigt väljer ut n objekt från en population om N objekt varav r objekt är OK och $N-r$ objekt är ej OK och X är antalet OK objekt man erhåller så är X enligt sats 1.4 och klassiska sannolikhetsformeln hypergeometrisk fördelad.

Det är intuitivt klart att $\mathbf{E}(X) = n\,(r/N)$ och $\mathbf{Var}(X)$ anges i boken.

<u>Simulering av diskret stokastisk variabel</u>. En <u>observation</u> av en stokastisk variabel X är värdet X får då man utför det slumpmässiga experimentet en gång och använder utfallet $\omega \in S$ till räkna ut den stokastiska variabelns värde $X(\omega)$ enligt definition 3.1.

Hur gör man en observation av en diskret stokastisk variabel X med möjliga värden $x_1, x_2, x_3 \ldots$ med respektive sannolikheter $f_X(x_1), f_X(x_2), f_X(x_3), \ldots$ i dator? Svar:

Sats 3.18. (Tabellmetoden) Tag ett slumptal ξ mellan 0 och 1 och sätt

$$X = \begin{cases} x_1 \text{ om} & 0 \le \xi < f(x_1) \\ x_2 \text{ om} & f(x_1) \le \xi < f(x_1) + f(x_2) \\ x_3 \text{ om} & f(x_1) + f(x_2) \le \xi < f(x_1) + f(x_2) + f(x_3) \\ \quad\vdots & \qquad\vdots \end{cases}.$$

Bevis. Sannolikheten att $\sum_{i=1}^{k-1} f_X(x_i) \le \xi < \sum_{i=1}^{k} f_X(x_i)$ är $f_X(x_k)$. □

Ovan är ξ egentligen en sk. kontinuerlig stokastisk variabel som är likformigt fördelad över intervallet $[0,1]$: detta läres ut i nästa kapitel 4.