

# Løsningsforslag for andre obligatoriske oppgave i

STK1100

Våren 2007

Av Ingunn Fride Tvette (ift@math.uio.no) og

Ørnulf Borgan (borgan@math.uio.no).

NB! Feil kan forekomme.

NB! Send gjerne en mail hvis du finner en feil!

## Oppgave 1

De stokastiske variablene  $X$  og  $Y$  har simultan sannsynlighetstetthet

$$f(x, y) = \begin{cases} k(x - y) & \text{for } 0 \leq y \leq x \leq 1 \\ 0 & \text{ellers} \end{cases}$$

a) Vi vet generelt at en tetthet skal integrere seg opp til 1.

Vi benytter det for å finne  $k$ .

$$\int_0^1 \int_0^x k(x-y) dy dx = k \int_0^1 [xy - \frac{1}{2}y^2]_0^x dx = k \int_0^1 (x^2 - \frac{1}{2}x^2) dx = k \int_0^1 \frac{1}{2}x^2 dx = k \frac{1}{6}.$$

Dette skal være lik 1, altså må  $k = 6$ .

$$\text{b) } P(2Y \leq X) = 6 \int_0^1 \int_0^{\frac{x}{2}} (x-y) dy dx = 6 \int_0^1 [xy - \frac{1}{2}y^2]_0^{\frac{x}{2}} dx = 6 \int_0^1 \frac{3}{8}x^2 dx = \frac{3}{4}.$$

c) For  $0 \leq x \leq 1$ :

$$f_X(x) = \int_{-\infty}^{\infty} f(x, y) dy = 6 \int_0^x (x-y) dy = 6 [xy - \frac{1}{2}y^2]_0^x = 6(x^2 - \frac{1}{2}x^2) = 3x^2.$$

Dette gir

$$f_X(x) = \begin{cases} 3x^2 & \text{for } 0 \leq x \leq 1 \\ 0 & \text{for ellers} \end{cases}$$

d) For  $0 \leq y \leq 1$ :

$$\begin{aligned} f_Y(y) &= \int_{-\infty}^{\infty} f(x, y) dx = 6 \int_y^1 (x-y) dy = 6 [\frac{1}{2}x^2 - yx]_y^1 \\ &= 6(\frac{1}{2} - y - \frac{1}{2}y^2 + y^2) = 3 - 6y + 3y^2. \end{aligned}$$

Dette gir

$$f_Y(y) = \begin{cases} 3 - 6y + 3y^2 & \text{for } 0 \leq y \leq 1 \\ 0 & \text{ellers} \end{cases}$$

e)  $X$  og  $Y$  er uavhengige så sant vi har at  $f(x, y) = f(x)f(y)$  for alle  $x$  og  $y$ . Nå har vi for  $0 \leq y \leq x \leq 1$  at

$$f_X(x)f_Y(y) = 3x^2(3 - 6y + 3y^2) = 9x^2 + 18x^2y + 9x^2y^2 \neq 6(x - y) = f(x, y).$$

Altså er  $X$  og  $Y$  ikke uavhengige.

## Oppgave 2

$X_1, X_2, \dots, X_n$  er uavhengige og identiske fordelte stokastiske variabler. Vi lar  $\mu = E(X_i)$  og  $\sigma^2 = Var(X_i)$ .

Videre har vi at  $\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i$  og at det standardiserte gjennomsnittet er gitt ved  $Z_n = \frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}} = \sqrt{n} \frac{\bar{X}_n - \mu}{\sigma}$ .

a) Vi skal vise at  $E(\bar{X}_n) = \mu$  og  $Var(\bar{X}_n) = \frac{\sigma^2}{n}$

$$\begin{aligned} E(\bar{X}_n) &= E\left(\frac{1}{n} \sum_{i=1}^n X_i\right) = \frac{1}{n} \sum_{i=1}^n E(X_i) = \frac{1}{n} \sum_{i=1}^n \mu = \frac{1}{n} n\mu = \mu. \\ Var(\bar{X}_n) &= Var\left(\frac{1}{n} \sum_{i=1}^n X_i\right) = \frac{1}{n^2} Var\left(\sum_{i=1}^n X_i\right) = \frac{1}{n^2} \sum_{i=1}^n Var(X_i) \\ &= \frac{1}{n^2} n\sigma^2 = \frac{\sigma^2}{n}. \end{aligned}$$

$$E(Z_n) = E\left(\frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}}\right) = \frac{1}{\sigma/\sqrt{n}} E(\bar{X}_n - \mu) = \frac{1}{\sigma/\sqrt{n}} (E(\bar{X}_n) - \mu) = \frac{1}{\sigma/\sqrt{n}} (\mu - \mu) = 0.$$

$$Var(Z_n) = Var\left(\frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}}\right) = \frac{1}{(\sigma/\sqrt{n})^2} Var(\bar{X}_n) = \frac{1}{(\sigma/\sqrt{n})^2} \frac{\sigma^2}{n} = 1.$$

b) Vi skal finne forventning, varians og standardavvik for den uniforme-, eksponential- og Bernoulli fordelingen, gitt ved hhv:

$$\begin{aligned} f(x) &= \begin{cases} 1 & \text{for } 0 \leq x \leq 1 \\ 0 & \text{ellers} \end{cases} \\ f(x) &= \begin{cases} e^{-x} & \text{for } x \geq 0 \\ 0 & \text{for } x < 0 \end{cases} \\ p(x) &= \frac{1}{2} \quad \text{for } x = 0, 1 \end{aligned}$$

Generelt har vi at

$$E(X^k) = \begin{cases} \int_{-\infty}^{\infty} x^k f(x) dx & \text{hvis } X \text{ kontinuerlig} \\ \sum_{-\infty}^{\infty} x^k p(x) & \text{hvis } X \text{ diskret} \end{cases}$$

$$\text{og } \text{Var}(X) = E(X^2) - (EX)^2.$$

1) Den uniforme fordeling:

$$\mu = E(X_i) = \int_0^1 x dx = [\frac{1}{2}x^2]_0^1 = \frac{1}{2}$$

$$E(X_i^2) = \int_0^1 x^2 dx = [\frac{1}{3}x^3]_0^1 = \frac{1}{3}$$

$$\sigma^2 = \text{Var}(X_i) = \frac{1}{3} - \left(\frac{1}{2}\right)^2 = \frac{1}{12}$$

$$\sigma = \sqrt{\text{Var}(X_i)} = \frac{1}{2\sqrt{3}}.$$

2) Eksponentialfordelingen:

$$\mu = E(X_i) = \int_0^\infty x e^{-x} dx = \Gamma(2) = 1 \cdot \Gamma(1) = 1$$

$$E(X_i^2) = \int_0^\infty x^2 e^{-x} dx = \Gamma(3) = 2 \cdot \Gamma(2) = 2$$

$$\sigma^2 = \text{Var}(X_i) = 2 - 1^2 = 1$$

$$\sigma = \sqrt{\text{Var}(X_i)} = 1.$$

3) Bernoulli fordeling

$$\mu = E(X_i) = 0 \cdot \frac{1}{2} + 1 \cdot \frac{1}{2} = \frac{1}{2}$$

$$E(X_i^2) = 0^2 \cdot \frac{1}{2} + 1^2 \cdot \frac{1}{2} = \frac{1}{2}$$

$$\sigma^2 = \text{Var}(X_i) = \frac{1}{2} - \left(\frac{1}{2}\right)^2 = \frac{1}{2} - \frac{1}{4} = \frac{1}{4}$$

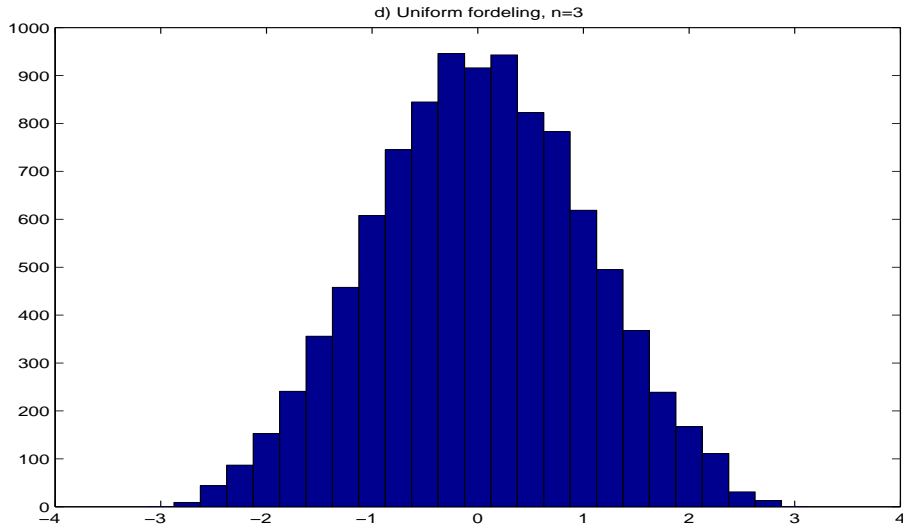
$$\sigma = \sqrt{\text{Var}(X_i)} = \frac{1}{2}$$

c) I notatet om “Kontinuerlige stokastiske variabler” (som fins under supplerende materiale på kursets hjemmeside) er det forklart at når vi har mange observasjoner av en kontinuerlig stokastisk variabel, så vil et normert histogram av observasjonene ligge nær sannsynlighetstettheten til den stokastiske variabelen. Når vi trekker fra den uniforme fordelingen eller fra eksponentialfordelingen, vil  $Z_n$  være en kontinuerlig stokastisk variabel. Derfor vil et histogram av 10000 observasjoner av  $Z_n$  ligge nær sannsynlighetstettheten til  $Z_n$ . Når vi trekker fra Bernoulli fordelingen, vil  $Z_n$  være en diskret stokastisk variabel. På lignende måte som i det kontinuerlige tilfellet kan en da argumentere for at et histogram av 10000 observasjoner av  $Z_n$  vil ligge nær punktsannsynligheten til  $Z_n$ .<sup>1</sup>

---

<sup>1</sup>Det at  $Z_n$  er diskret fordelt i Bernoulli tilfellet skaper noen komplikasjoner, som det ikke

d) Vi utfører kommandoen gitt i oppgaveteksten og får histogrammet under:



Figur 1: Histogrammer over  $Z_3$  verdier, der hver  $Z_3$  verdi er basert på 3 trukne  $X$  verdier fra den uniforme fordeling.

Histogrammet er noenlunde symmetrisk fordelt om 0, og ligner ganske mye på en normalfordeling (men se punkt f nedenfor). [At histogrammet er noenlunde symmetrisk er ikke overraskende, for den uniforme fordelingen er symmetrisk om sin forventningsverdi.] Det at det er to toppe på histogrammet i vår simulering, er det ikke noen grunn til å legge vekt på.<sup>2</sup>

forventes at studentene går inn på. Når vi trekker  $n$  observasjoner fra Bernoulli fordelingen, vil  $\sum_{i=1}^n X_i$  kunne anta verdiene  $k = 0, 1, \dots, n$ . Vi kan derfor illustrere punktsannsynligheten til  $\sum_{i=1}^n X_i$  som et "sannsynlighetshistogram" med  $n + 1$  "stolper" som hver har bredde lik 1 og punkt i  $0, 1, \dots, n$ , og der høyden av stolpene (og dermed også arealet) er lik sannsynlighetene for at  $\sum_{i=1}^n X_i$  antar de ulike verdiene.

Men hvordan kan vi illustrere punktsannsynlighetene til  $Z_n = \sqrt{n}(\bar{X}_n - 1/2)/(1/2) = \frac{2}{\sqrt{n}}(\sum_{i=1}^n X_i/n - \frac{n}{2})$ ? Vi merker oss at  $Z_n$  er en diskret fordelt stokastisk variabel som antar verdien  $\frac{2}{\sqrt{n}}(k - \frac{n}{2})$  med samme sannsynlighetene som  $\sum_{i=1}^n X_i$  antar verdien  $k$ ;  $k = 0, 1, \dots, n$ . Når vi skal illustrere punktsannsynligheten til  $Z_n$  kan vi tegne stolper med bredde  $2/\sqrt{n}$  som har midtpunkt i  $\frac{2}{\sqrt{n}}(k - \frac{n}{2})$  for  $k = 0, 1, \dots, n$ . Det er to mulige valg for høydene av stolpene. Vi kan la dem være lik sannsynlighetene for de ulike verdiene av  $Z_n$ , eller vi kan la høydene være lik  $\frac{\sqrt{n}}{2}$  ganger disse sannsynlighetene. I det siste tilfellet vil arealet av hver stolpe være lik den tilsvarende sannsynligheten. Hvis vi velger dette siste alternativet, vil et histogram av 10000 observasjoner av  $Z_n$  ligge nær det tilsvarende sannsynlighetshistogrammet for punktsannsynligheten til  $Z_n$  i Bernoulli tilfellet.

<sup>2</sup>Når vi simulerer vil arealet av en "stolpe" i histogrammet være lik den relative frekvensen

e/f) Vi sammenligner den empiriske fordelingen til  $Z_n$  med standard normalfordelingen. Vi beregner den relative frekvensen av verdier av  $Z_n$  i de oppgitte intervallene og sammenligner med sannsynligheten for at en standard normalfordelt variabel vil falle i det samme intervallet. Resultatene av dette er presentert i tabellen nedenfor:

Int.	$(-\infty, -2.5)$	$[-2.5, -2.0)$	$[-2.0, -1.5)$	$[-1.5, -1.0)$	$[-1.0, -0.5)$	$[-0.5, 0.0)$
Std norm.	0.0062	0.0166	0.0440	0.0919	0.1498	0.1915
$Z_3$	0.0028	0.0172	0.0508	0.0905	0.1482	0.1862
Intervall	$[0.0, 0.5)$	$[0.5, 1.0)$	$[1.0, 1.5)$	$[1.5, 2.0)$	$[2.0, 2.5)$	$[2.5, \infty)$
Std norm.	0.1915	0.1498	0.0919	0.0440	0.0166	0.0062
$Z_3$	0.1823	0.1479	0.1012	0.0502	0.0203	0.0024

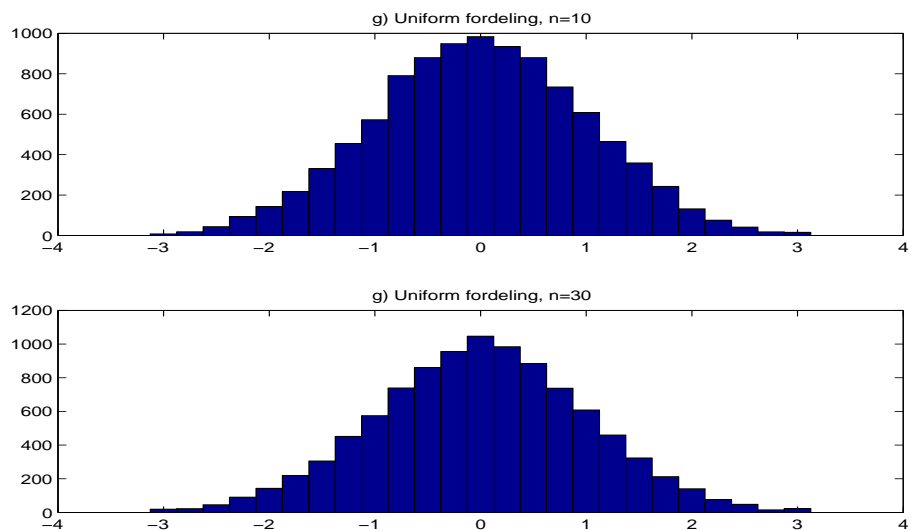
Vi ser at det er nokså god overensstemmelse mellom de relative frekvensene av verdier av  $Z_3$  i de oppgitte intervallene og sannsynlighetene for at en standard normalfordelt variabel vil falle i de samme intervallene.

Hvis vi ser nøyere etter, merker vi oss at det ser ut til at det er litt mindre sannsynlig at  $Z_3$  vil anta ekstreme verdier"[dvs. i intervallene  $(-\infty, -2.5)$  og  $[2.5, \infty)$ ] enn tilfellet er for en standardnormalfordelt variabel.

g) Vi gjør det samme som i d) og f) for den uniforme fordeling når  $n = 10$  og  $n = 30$ .

---

av  $Z_n$ -verdier i det aktuelle intervallet. Og denne relative frekvensen er bare *tilnærmet* lik sannsynligheten for at  $Z_n$  vil få en verdi i intervallet. Så hvis vi simulerer på nytt, vil høyden av stolpene forandre seg noe, og spesielt kan de to toppene forsvinne.



Figur 2: Histogrammer over  $Z_n$  verdier, der hver  $Z_n$  verdi er basert på hhv 10 og 30 trukne  $X$  verdier fra den uniforme fordeling.

Både for  $n = 10$  og for  $n = 30$  er histogrammene noenlunde symmetriske om 0, og de ligner på en normalfordeling (se også kommentarene til tabellen nedenfor).

Litt generell teori:

La oss anta at vi har en  $X$  der  $X \sim N(\mu, \sigma^2)$ .

Vi har at dens pdf har sitt maksimum ved  $X = \mu$ , og vendepunkt (der kurven endrer seg fra konkav til konveks) ved  $\mu \pm \sigma$ .

Videre vet vi for en normalfordeling at

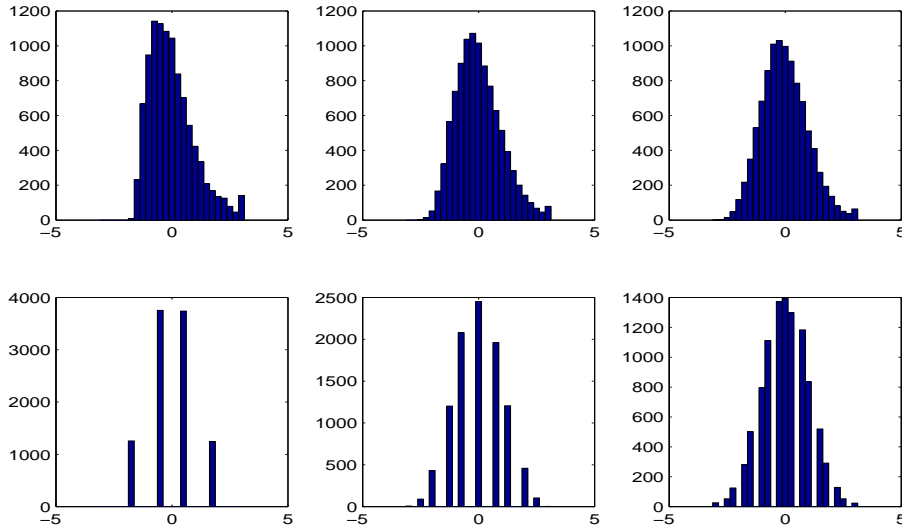
$$\begin{aligned} P(|X - \mu| \leq \sigma) &= P(|Z| \leq 1) = 0.6826, \\ P(|X - \mu| \leq 2\sigma) &= P(|Z| \leq 2) = 0.9544 \text{ og} \\ P(|X - \mu| \leq 3\sigma) &= P(|Z| \leq 3) = 0.9974 \end{aligned}$$

Vi kan sjekke om dette er tilfellet her (prøv selv!).

Int.	$(-\infty, -2.5)$	$[-2.5, -2.0)$	$[-2.0, -1.5)$	$[-1.5, -1.0)$	$[-1.0, -0.5)$	$[-0.5, 0.0)$
Std norm.	0.0062	0.0166	0.0440	0.0919	0.1498	0.1915
$Z_{10}$	0.0045	0.0187	0.0438	0.0902	0.1550	0.1886
$Z_{30}$	0.0064	0.0188	0.0435	0.0874	0.1479	0.1921
Int.	$[0.0, 0.5)$	$[0.5, 1.0)$	$[1.0, 1.5)$	$[1.5, 2.0)$	$[2.0, 2.5)$	$[2.5, \infty)$
Std norm.	0.1915	0.1498	0.0919	0.0440	0.0166	0.0062
$Z_{10}$	0.1881	0.1470	0.0972	0.0460	0.0163	0.0046
$Z_{30}$	0.1954	0.1496	0.0905	0.0459	0.0165	0.0060

Vi ser igjen en nokså god overensstemmelse mellom de relative frekvensene av verdier av  $Z_{10}$  og  $Z_{30}$  i de oppgitte intervallene og sannsynlighetene for at en standard normalfordeling variabel vil falle i de samme intervallene. For  $n = 10$  ser vi imidlertid fortsatt noe av det samme som vi fant for  $n = 3$ . Det er litt mindre sannsynlig at  $Z_{10}$  vil anta verdier i intervallene  $(-\infty, -2.5)$  og  $[2.5, \infty)$  enn tilfellet er for en standardnormalfordelt variabel. For  $n = 30$  ser vi ikke lenger dette fenomenet.

h)/i) Vi gjør tilsvarende simuleringer for eksponentialfordelingen og Bernoulli-fordelingen, med  $n = 3$ ,  $n = 10$  og  $n = 30$ .



Figur 3: Histogrammer over  $Z_n$  verdier, der hver  $Z_n$  verdi er basert på hhv 3, 10 og 30 trukne  $X$  verdier fra hhv eksponential fordelingen (øverst) og Bernoulli fordelingen (nederst).

Vi ser først på de 3 øverste histogrammene, der den eksponentielle fordelingen ligger til grunn:

Når  $Z_n$  kun er basert på 3 trukne  $X$  verdier ser histogrammet temmelig skjevt ut, og ligner ikke mye på en normal fordeling. Jo større  $n$  blir jo mer symmetrisk blir histogrammet om 0, og for  $n = 30$ , dvs for histogrammet øverst til høyre, ligner histogrammet på en normalfordeling (se også kommentar til tabellen nedenfor).

Vi ser på de 3 nederste histogrammene, der Bernoulli fordelingen ligger til grunn:

Vi ser at alle histogrammene er noenlunde symmetriske om 0. [Bernoulli fordelingen med  $p = 1/2$  er jo symmetrisk om sin forventningsverdi, i motsetning til eksponentialfordelingen som er skjev.] Men vi ser også at  $Z_n$  er en diskret stokastisk variabel når vi trekker fra Bernoulli fordelingen. Dette er spesielt tydelig når  $n = 3$  og  $n = 10$ . Når  $n = 30$  syns det fortsatt at vi har en diskret fordeling, men siden antall mulige verdier nå er nokså stort, er den “diskrete naturen” av fordelingen



ikke lenger så tydelig, og histogrammet ligner nokså mye på en normalfordeling.<sup>3</sup>

Den eksponentielle fordelingen:

Int.	$(-\infty, -2.5)$	$[-2.5, -2.0)$	$[-2.0, -1.5)$	$[-1.5, -1.0)$	$[-1.0, -0.5)$	$[-0.5, 0.0)$
Std norm.	0.0062	0.0166	0.0440	0.0919	0.1498	0.1915
$Z_3$	0	0	0.0065	0.1302	0.2199	0.2204
$Z_{10}$	0	0.0033	0.0345	0.1097	0.1776	0.2173
$Z_{30}$	0.0006	0.0108	0.0446	0.1045	0.1718	0.2058
Int.	$[0.0, 0.5)$	$[0.5, 1.0)$	$[1.0, 1.5)$	$[1.5, 2.0)$	$[2.0, 2.5)$	$[2.5, \infty)$
Std norm.	0.1915	0.1498	0.0919	0.0440	0.0166	0.0062
$Z_3$	0.1700	0.1086	0.0648	0.0337	0.0225	0.0234
$Z_{10}$	0.1753	0.1255	0.0797	0.0404	0.0212	0.0155
$Z_{30}$	0.1810	0.1316	0.0798	0.0398	0.0172	0.0125

For  $n = 3$  avviker de relative frekvensene klart fra de tilsvarende sannsynlighetene for standardnormalfordelingen. Vi ser også av de relative frekvensene at fordelingen er temmelig skjev. For eksempel er de relative frekvensene lik 0 for de to intervallene  $(-\infty, -2.5)$  og  $[-2.5, -2.0)$ , mens de er litt over 2% for de to “symmetrisk plasserte” intervallene  $[2.0, 2.5)$  og  $[2.5, \infty)$ . Også for andre “symmetrisk plasserte” intervaller [f.eks.  $[-1.5, -1.0)$  og  $[1.0, 1.5)$ ] er det klar forskjell på de relative frekvensene når  $n = 3$ .

Også når  $n = 10$  avviker de relative frekvensene en god del fra de tilsvarende sannsynlighetene for standardnormalfordelingen, og det er forskjell på de relative frekvensene for “symmetrisk plasserte” intervaller. Men forskjellen er klart mindre enn når  $n = 3$ .

For  $n = 30$  er histogrammet nokså symmetrisk, og de relative frekvensene er nokså nær de tilsvarende sannsynlighetene for standardnormalfordelingen. Men det er fortsatt tydelige avvik fra symmetri og normalfordelingen; se f.eks. intervallene  $[-1.0, -0.5)$  og  $[0.5, 1.0)$ .

Benoulli fordelingen:

Int.	$(-\infty, -2.5)$	$[-2.5, -2.0)$	$[-2.0, -1.5)$	$[-1.5, -1.0)$	$[-1.0, -0.5)$	$[-0.5, 0.0)$
Std norm.	0.0062	0.0166	0.0440	0.0919	0.1498	0.1915
$Z_3$	0	0	0.1257	0	0.3755	0
$Z_{10}$	0.0099	0	0.0432	0.1202	0.2080	0
$Z_{30}$	.0079	0.0125	0.0282	0.1299	0.1112	0.1372
Int.	$[0.0, 0.5)$	$[0.5, 1.0)$	$[1.0, 1.5)$	$[1.5, 2.0)$	$[2.0, 2.5)$	$[2.5, \infty)$
Std norm.	0.1915	0.1498	0.0919	0.0440	0.0166	0.0062
$Z_3$	0	0.3742	0	0.1246	0	0
$Z_{10}$	0.2452	0.1961	0.1207	0.0459	0	0.0108
$Z_{30}$	0.2692	0.1184	0.1357	0.0292	0.0129	0.0077

<sup>3</sup>I denne oppgaven har vi for enkelhets skyld valgt å tegne alle histogrammene med klassebredde lik 0.25. For Bernoulli tilfellet hadde det vært bedre å la klassebredden og plasseringen av intervallene avhenge av  $n$  slik det er beskrevet i fotnoten til punkt c. Da ville vi unngått å få “huller” i histogrammene, og skalaen langs  $y$ -aksen hadde blitt den samme for alle  $n$ . (Slik histogrammene blir lagd i denne oppgaven, vil ikke arealet av alle søylene til samme bli lik 1 for Bernoulli tilfellet.)

Særlig for  $n = 3$ , men også for  $n = 10$  ser vi tydelig den “diskrete naturen” til fordelingen til  $Z_n$ . Den gjør seg gjeldende ved at det er umulig for  $Z_n$  å få verdier i noen av intervallene. De relative frekvensene ligger da også nokså langt fra de tilsvarende sannsynlighetene for standardnormalfordelingen.

Når  $n = 30$  er de relative frekvensene ganske nær de tilsvarende tilsvarende sannsynlighetene for standardnormalfordelingen. De største avvikene har vi for intervallene  $[-0.5, 0)$  og  $[0, 0.5)$ . Det skyldes at  $Z_{30}$  har en mulig verdi (nemlig  $-0.365$ ) i det første av disse intervallene og to mulige verdier (nemlig  $0$  og  $0.365$ ) i det andre. Hvis vi hadde fordelt sannsynligheten for verdien  $0$  likt på de to intervallene, ville standardnormalfordelingen ha passert ganske godt også for disse.

j) Vi sammenligner resultatene.

I punktene d,f og g fant vi at standardnormalfordelingen gir en rimelig god tilnærming til fordelingen for  $Z_n$  alt når  $n = 3$  når vi trekker fra en uniform fordeling. Og for  $n = 10$  og  $n = 30$  er tilnærmingen temmelig god. At vi får en så god tilnærming alt for små verdier av  $n$  kommer av at den uniforme fordelingen er (i) kontinuerlig og (ii) symmetrisk [og (iii) har “lette haler”].

I punkt g fant vi at standardnormalfordelingen ga en veldig dårlig tilnærming til fordelingen for  $Z_n$  når vi trekker fra eksponentialfordelingen og  $n = 3$ . Tilnærmingen er heller ikke god når  $n = 10$ , men for  $n = 30$  begynner det å hjelpe. Grunnen til at vi må ha større verdi av  $n$  for eksponentialfordelingen enn for den uniforme fordelingen er at eksponentialfordelingen er skjev [og at den har en tung hale”].

I punkt i fant vi at standardnormalfordelingen ga en veldig dårlig tilnærming til fordelingen for  $Z_n$  når vi trekker fra Bernoulli fordelingen og  $n = 3$ . Tilnærmingen er også nokså dårlig når  $n = 10$ . For  $n = 30$  begynner det imidlertid å hjelpe. Grunnen til at vi må ha større verdi av  $n$  for Bernoulli fordelingen enn for den uniforme fordelingen er at Bernoulli fordelingen er diskret.

Sentralgrensesetningen i avsnitt 5.3 i læreboka forteller oss at fordelingen til  $Z_n$  vil nærme seg standardnormalfordelingen når  $n$  blir stor. I denne oppgaven har vi sett at hvor fort fordelingen til  $Z_n$  vil nærme seg standardnormalfordelingen avhenger av fordelingen til  $X_i$ -ene. Det går fortere for kontinuerlige fordelinger enn for diskrete fordelinger, og det går fortere for symmetriske fordelinger enn for skjeve fordelinger [og det går fortere for fordelinger med “lette haler” enn for fordelinger med “tunge haler”].