

LØSNING: Oppgavesett nr. 3

“MAT110 Statistikk I”, 2019

Oppgave 1: (Sport Management - binomisk fordeling , forventning , varians)

a) Forsøksserien med å prøve å score mål i en fotballkamp med $n = 10$ målsjanser har følgende **egenskaper**:

1. Hver målsjanse har kun 2 mulige utfall, mål (suksess) eller ikke mål (fiasko).
2. I oppgaven antas det at det er **samme sannsynlighet** p ($= 0.20$) for å score mål i alle n ($= 10$) forsøkene.
3. Alle forsøk er **uavhengige**.
4. Vi gjennomfører et bestemt antall forsøk, $n = 10$ i dette tilfellet.

Forsøksserien oppfyller dermed kravene til en **binomisk** forsøksserie. Den stokastiske variablene X som beskriver antallet suksesser i denne forsøksserien er derfor **binomisk** fordelt.

b) i) Tolking:

$$\underline{\underline{E[X] = \text{forventet antall mål som scores av "Tuftes IL" i løpet av tilfeldig valgt kamp}}}$$
(1)

ii) Siden $X \sim \text{Bin}[n = 10, p = 0.20]$ så er forventningen:

$$\underline{\underline{E[X] \stackrel{\text{Bin.}}{=} np = 10 \cdot 0.2 = 2}}$$
(2)

c) i) Tolking:

$$\underline{\underline{Var[X] = \text{forventet variasjon/spredning av antall mål som scores av} \\ \text{“Tufte IL” i løpet av tilfeldig valgt kamp}}}} \quad (3)$$

ii) Siden $X \sim \text{Bin}[n = 10, p = 0.20]$ så er variansen:

$$\underline{\underline{Var[X] \stackrel{\text{Bin.}}{=} np(1-p) = 10 \cdot 0.2 \cdot (1 - 0.2) = \underline{\underline{1.6}}}}} \quad (4)$$

d) i) Sannsynligheten for at laget scorer minst 1 mål i en kamp:

$$\underline{\underline{P(X \geq 1)}} = 1 - P(X \leq 0) \quad (5)$$

$$= 1 - P(X = 0) \quad (6)$$

$$= 1 - \binom{n}{0} p^0 (1-p)^{n-0} = 1 - \binom{10}{0} 0.2^0 (1-0.2)^{10-0} = \underline{\underline{0.8926}} \quad (7)$$

ii) Sannsynligheten for at laget scorer minst 2 mål i en kamp:

$$\underline{\underline{P(X \geq 2)}} = 1 - P(X = 0) - P(X = 1) \quad (8)$$

$$= 1 - \binom{n}{0} p^0 (1-p)^{n-0} - \binom{n}{1} p^1 (1-p)^{n-1} \quad (9)$$

$$= 1 - \binom{10}{0} 0.2^0 (1-0.2)^{10-0} - \binom{10}{1} 0.2^1 (1-0.2)^{10-1} \quad (10)$$

$$= 1 - 0.1074 - 0.2684 \quad (11)$$

$$= \underline{\underline{0.6242}} \quad (12)$$

e) Utfylling tabell:

	$E[X]$	$\sigma[X]$
Historiske data	2	1.26
Strategi A (øke "p")	3	1.45
Strategi B (øke "n")	3	1.55

Figur 1: $E[X]$ og $\sigma[X]$.

- f) Fra tabell 1 ser vi at strategi A og B gir samme forventede antall scorede mål per kamp. I tillegg er standardavvikene også velge like, men strategi A har litt mindre standardavvik $\sigma[X]$. Det betyr at det er litt mindre usikkerhet forbundet med strategi A.

Konklusjonen er derfor at:

Ingen dramatisk forskjell på de to strategiene, men strategi A er litt bedre siden den har mindre usikkerhet.



Oppgave 2: (petroleumslogistikk - Poisson fordeling , sum av stokastiske variabler)

a) Siden $X \sim \text{Poi}[X]$ og 1 “hendelse” per uke:

$$\underline{P(X = 1)} = \frac{\lambda^1}{1!} e^{-\lambda} \quad (13)$$

$$\stackrel{\lambda=0.6}{=} \frac{0.6^1}{1!} e^{-0.6} \stackrel{\text{kalkis}}{=} \underline{0.3293} \quad (14)$$

Sannsynligheten for at det skjer 1 “hendelse” per uke er 32.93%.

b) Siden $X \sim \text{Poi}[X]$ og mer enn 1 “hendelse” per uke:

$$\underline{P(X > 1)} = 1 - P(X \leq 1) \quad (15)$$

$$= 1 - \left(P(X = 0) + P(X = 1) \right) \quad (16)$$

$$= 1 - \frac{\lambda^0}{0!} e^{-\lambda} - \frac{\lambda^1}{1!} e^{-\lambda} \quad (17)$$

$$\stackrel{\lambda=0.6}{=} 1 - \frac{0.6^0}{0!} e^{-0.6} - \frac{0.6^1}{1!} e^{-0.6} \quad (18)$$

$$\stackrel{\text{kalkis}}{=} 1 - 0.5488 - 0.3293 \quad (19)$$

$$= \underline{0.1219} \quad (20)$$

Sannsynligheten for at det skjer mer enn 1 “hendelse” per uke er 12.19%.

- c) Forventet **antall** “hendelser” per 2 uker:

$$\underline{\underline{E[Y]}} = E[X_1 + X_2] \quad (21)$$

$$\stackrel{\text{alltid}}{=} E[X_1] + E[X_2] = \overbrace{\lambda + \lambda}^{2 \text{ stk.}} \quad (22)$$

$$= 2\lambda = 2 \cdot 0.6 = \underline{\underline{1.2}} \quad (23)$$

NB: Overgangen i lign.(60) til (61) gjelder **alltid**. Uansett om de stokastiske variablene X_i er uavhengige eller ikke. ¹

- d) En sum av Poisson fordelinger $X_1 + X_2$ er også Poisson fordelt, dvs. Y er Poisson fordelt.

Den stokastiske variabelen Y er da Poisson fordelt med forventning $\lambda_Y = E[Y] = 1.2$.

Sannsynligheten for at det skjer *mer enn* 1 “hendelse” i løpet er av periode på 2 uker er da:

$$\underline{\underline{P(Y > 1)}} = 1 - P(Y \leq 1) \quad (24)$$

$$= 1 - \left(P(Y = 0) + P(Y = 1) \right) \quad (25)$$

$$= 1 - \frac{\lambda_Y^0}{0!} e^{-\lambda_Y} - \frac{\lambda_Y^1}{1!} e^{-\lambda_Y} \quad (26)$$

$$\stackrel{\lambda_Y=1.2}{=} 1 - \frac{1.2^0}{0!} e^{-1.2} - \frac{1.2^1}{1!} e^{-1.2} \quad (27)$$

$$\stackrel{\text{kalkis}}{=} 1 - 0.3012 - 0.3614 \quad (28)$$

$$= \underline{\underline{0.3374}} \quad (29)$$

¹Se ”regneregler” angående forventningsverdier i kompendiet.

e) Generelle addisjonssetning:

$$P(X_1 > 1 \text{ eller } X_2 > 1) = \underbrace{P(X_1 > 1) + P(X_2 > 1)}_{= 2P(X > 1)} - \overbrace{P(X_1 > 1, X_2 > 1)}^{\text{og-sannsynlighet}} = P(X > 1)^2 \quad (30)$$

Sannsynligheten $P(X > 1)$ for en gitt uke er den samme uansett uke, dvs. $P(X_1 > 1) = P(X_2 > 1) = P(X > 1)$.

Siden X_i er uavhenige så bruker vi den spesielle multiplikasjonssetningen på og-sannsynligheten.

Bruker at $P(X > 1) = 0.1219$ fra oppgave **2b**:

$$\underline{P(X_1 > 1 \text{ eller } X_2 > 1)} = 2P(X > 1) - P(X > 1)^2 \quad (31)$$

$$= 2 \cdot 0.1219 - 0.1219^2 = \underline{0.2289} \quad (32)$$

Sannsynligheten for at det skjer mer enn én hendelse i uke 1 ELLER at det skjer mer enn én hendelse i uke 2, er 0.2298.

f) Siden

$$P(Y > 1) = P(X_1 > 1 \text{ eller } X_2 > 1) + P(X_1 = 1, X_2 = 1) \quad (33)$$

så er:

$$P(Y > 1) - P(X_1 > 1 \text{ eller } X_2 > 1) = \underbrace{P(X_1 = 1, X_2 = 1)}_{\geq 0} \quad (34)$$

Siden $P(X_1 = 1, X_2 = 1) \geq 0$ så er også alltid $P(Y > 1) \geq P(X_1 > 1 \text{ eller } X_2 > 1)$.

■

Oppgave 3: (logistikk og økonomi - binomisk , sannsynlighet , forventning , varians)

a) “Forsøksserien” med oppmøte til en flyavgang har følgende egenskaper:

1. For hver passasjer er det kun 2 mulige utfall, oppmøte (suksess) eller ikke oppmøte (fiasko).
2. I oppgaven antas samme sannsynlighet p for oppmøte for alle n billettkjøperne.
3. I oppgaven antas det at alle billettkjøperne møter opp uavhengig av hverandre.
4. Det gjennomføres et bestemt antall forsøk, dvs. et bestemt antall passasjerer i dette tilfellet.

Forsøksserien oppfyller dermed kravene til en **binomisk** forsøksserie. Den stokastiske variablene X som beskriver antallet suksesser i denne forsøksserien er derfor binomisk fordelt.

b) i) Forventning av $X \sim \text{Bin}[n, p]$:

$$\underline{\underline{E[X]}} = np = 120 \cdot 0.95 = \underline{\underline{114}} \quad (35)$$

ii) Tolkning:

$$\underline{\underline{E[X]}} = \text{forventet antall billettkjøpere som faktisk møter opp til sin flyavgang}$$

c) i) Variansen til $X \sim \text{Bin}[n, p]$:

$$\underline{\underline{Var[X]}} = np(1-p) = 120 \cdot 0.95 \cdot (1 - 0.95) = \underline{\underline{5.7}} \quad (36)$$

ii) Tolkning:

$$\underline{\underline{Var[X]}} = \text{forventet varians/usikkerhet i antall billettgjørere} \\ \underline{\underline{\text{som faktisk møter opp til sin flyavgang}}}$$

d) Forventet inntekt for SAS når $n = 120$:

$$\underline{\underline{E[I]}} = E[a X] = a \underbrace{E[X]}_{=np} = 800 \cdot np \text{ NOK} \quad (37)$$

$$= 800 \cdot 120 \cdot 0.95 \text{ NOK} = \underline{\underline{91\,200 \text{ NOK}}} \quad (38)$$

e) Forventet inntekt for SAS når $n = 123$:

$$\underline{\underline{E[I]}} = E[a X] = a \underbrace{E[X]}_{=np} = 800 \cdot np \text{ NOK} \quad (39)$$

$$= 800 \cdot 123 \cdot 0.95 \text{ NOK} = \underline{\underline{93\,480 \text{ NOK}}} \quad (40)$$

f) Sannsynligheten for at det møter opp flere passasjerer enn flyet har kapasitet til:

$$\underline{\underline{P(X \geq 121)}} = P(X = 121) + P(X = 122) + P(X = 123) \quad (41)$$

$$= \binom{n}{121} p^{121} (1-p)^{n-121} + \binom{n}{122} p^{122} (1-p)^{n-122} + \binom{n}{123} p^{123} (1-p)^{n-123}$$

$$= \binom{123}{121} 0.95^{121} (1-0.95)^{123-121} + \binom{123}{122} p^{122} (1-0.95)^{123-122} \quad (42)$$

$$+ \binom{123}{123} 0.95^{123} (1-0.95)^{123-123} \quad (43)$$

$$= 0.0378 + 0.0118 + 0.0018 \quad (44)$$

$$= \underline{\underline{0.0514}} \quad (\text{eksakt svar med 4 desimales nøyaktighet}) \quad (45)$$

Kommentar:

Det er meningen at denne oppgaven skal løses på måten som vist i lign.(41). Det er en eksakt løsning. Man trenger nemlig mellomregningene, dvs. lign.(44), for å løse oppgave **3g**.

I kapittel 4 skal vi lære om normalfordelingen.

Da kan også oppgave **3f** løses via en tilnærmet metode. Selv om det er meningen å løse denne oppgaven eksakt som vist ovenfor, så nevner vi likevel hvordan kan løse oppgaven med en tilnærmet metode: ²

Den tilnærmede metoden gir ikke et så godt svar som det eksakte, selvsagt. Men man får et svar som er i nærheten. Siden

$$n \cdot p(1-p) = 5.8 \gtrsim 5 \quad (46)$$

så er X *tilnærmet* en normalfordeling, $X \sim N[E[X], \sigma[X]]$, hvor $E[X] = np = 116.85$ og $Var[X] = np(1-p) = 5.8425$ ($n = 123$ og $p = 0.95$). Uten heltallskorreksjon får man da:

$$\underline{\underline{P(x \geq 121)}} = 1 - P(Z \leq 1.72) = 1 - 0.9573 = \underline{\underline{0.0427}} \quad (\text{tilnærmet svar}) \quad (47)$$

²Man behøver ikke å ha med denne kommentaren i innleveringen.

g) Tolkning:

$$\underline{\underline{E[C]}} = \text{forventet kostnad for SAS pga. overbooking} \quad (48)$$

h) Forventet kostnad til SAS pga. “overbooking”, når $n = 123$, altså det selger 3 flere billetter enn det er plass til på flyet:

$$\underline{\underline{E[C]}} = b \sum_{x=121}^{123} (x - 120) P(X = x) \quad (49)$$

$$= b \left((121 - 120)P(X = 121) + (122 - 120)P(X = 122) + (123 - 120)P(X = 123) \right)$$

$$= b \left(\underbrace{1 \cdot \underbrace{P(X = 121)}_{= 0.0378} + 2 \cdot \underbrace{P(X = 122)}_{= 0.0118} + 3 \cdot \underbrace{P(X = 123)}_{= 0.0018}}_{\text{numeriske resultat hentet fra oppgave 3f, lign. (44)}} \right) \quad (50)$$

$$= 5000 \cdot \left(1 \cdot 0.0378 + 2 \cdot 0.0118 + 3 \cdot 0.0018 \right) \text{ NOK} \quad (51)$$

$$= \underline{\underline{334 \text{ NOK}}} \quad (52)$$

i) Siden forventet billettinntekter ved overbooking er

$$\underline{\underline{E[I] - E[U]}} = \left(\overbrace{93\,480}^{\text{oppg. 2e}} - 334 \right) \text{ NOK} = \underline{\underline{93\,146 \text{ NOK}}} \quad (53)$$

er større enn forventet inntekt ved fullt fly, $\overbrace{91\,200}^{\text{oppg. 3d}} \text{ NOK}$
 så lønner det seg for SAS å overbooke.

■

Oppgave 4: (logistikk - Poisson fordeling , forventning , varians , sannsynligheter)

- a) Det er oppgitt i oppgaven at $X_i \sim \text{Poi}[\lambda]$.
Siden $X_G \sim \text{Poi}[\lambda]$ så finner man sannsynligheten for at det skjer **2 utrykninger** i løpet av en uke i Gjemnes på følgende måte:

$$\underline{\underline{P(X_G = 2)}} = \frac{\lambda_G^2}{2!} e^{-\lambda_G} \quad (54)$$

$$\lambda_G \equiv 1.9 \quad \frac{1.9^2}{2!} e^{-1.9} \stackrel{\text{kalkis}}{=} \underline{\underline{0.2700}} \quad (55)$$

- b) Det er oppgitt i oppgaven at $X_G \sim \text{Poi}[\lambda]$.
Sannsynligheten for at det skjer **mer enn 2** utrykninger i Gjemnes i løpet av en uke:

$$\underline{\underline{P(X_G > 2)}} = 1 - P(X_G \leq 2) \quad (56)$$

$$= 1 - \left(\overbrace{P(X_G = 0) + P(X_G = 1)}^{\text{oppgitt i oppgaven.}} + \overbrace{P(X_G = 2)}^{\text{fra oppg. 4a}} \right) \quad (57)$$

$$= 1 - 0.1496 - 0.2842 - 0.2700 \quad (58)$$

$$= \underline{\underline{0.2962}} \quad (59)$$

- c) Forventet **antall** akutte **utrykninger** i Eide, Fræna og Gjemnes til sammen per uke:

$$\underline{\underline{E[Y]}} = E[X_E + X_F + X_G] \quad (60)$$

$$\stackrel{\text{alltid}}{=} E[X_E] + E[X_F] + E[X_G] = \lambda_E + \lambda_F + \lambda_G \quad (61)$$

$$= 1.5 + 2.3 + 1.9 = \underline{\underline{5.7}} \quad (62)$$

NB: Overgangen i lign.(60) til (61) gjelder **alltid**. Uansett om de stokastiske variablene X_i er uavhengige eller ikke.

- d) En sum av Poisson fordelinger $X_E + X_F + X_G$ er også Poisson fordelt, dvs. Y er Poisson fordelt med forventning $\lambda_Y = E[Y] = 5.7$.

Sannsynligheten for at det skjer *mer enn* 2 akutte utrykningen i Eide, Fræna og Gjemnes tilsammen oer uke er da:

$$\underline{\underline{P(Y > 2)}} = 1 - P(Y \leq 2) \quad (63)$$

$$= 1 - \left(P(Y = 0) + P(Y = 1) + P(Y = 2) \right) \quad (64)$$

$$= 1 - \frac{\lambda_Y^0}{0!} e^{-\lambda_Y} - \frac{\lambda_Y^1}{1!} e^{-\lambda_Y} - \frac{\lambda_Y^2}{2!} e^{-\lambda_Y} \quad (65)$$

$$\stackrel{\lambda_Y=5.7}{=} 1 - \frac{5.7^0}{0!} e^{-5.7} - \frac{5.7^1}{1!} e^{-5.7} - \frac{5.7^2}{2!} e^{-5.7} \quad (66)$$

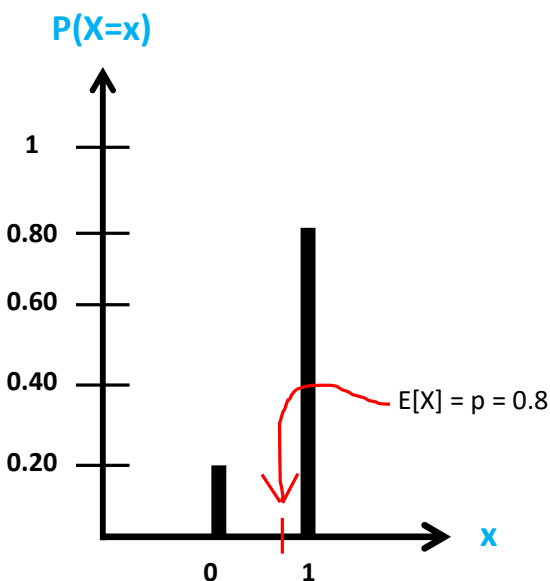
$$\stackrel{\text{kalkis}}{=} 1 - 0.0033 - 0.0191 - 0.0544 \quad (67)$$

$$= \underline{\underline{0.9232}} \quad (68)$$

■

Oppgave 5: (medisin - Bernoulli-fordelingen og kobling med binomial-fordelingen)

a) Med $p = 0.8$ så er sannsynlighetsfordelingen til $X \sim \text{Ber}[p]$ gitt ved: ³



Figur 2: Sannsynlighetsfordeling $P(X = x)$.

b) Bruker definisjonen av forventning:

$$\underline{\underline{E[X]}} \stackrel{\text{def.}}{=} \sum_{i=0}^1 x_i P(X = x_i) \quad (69)$$

$$= x_0 P(X = x_0) + x_1 P(X = x_1) \quad (70)$$

$$= 0 \cdot (1 - p) + 1 \cdot p = \underline{\underline{p}} \quad (71)$$

altså forventningsverdien til $X \sim \text{Ber}[p]$ er bare p .

³Du trenger ikke ha med $E[X] = p = 0.8$ i figur 2. Forventningen $E[X] = p = 0.8$ er bare tatt med for å illustrere resultatet fra oppgave **5b**.

c) Bruker definisjonen av varians:

$$\underline{\underline{Var[X]}} \stackrel{\text{def.}}{=} \sum_{i=0}^1 \left(x_i - E[X] \right)^2 P(X = x_i) \quad (72)$$

$$= \left(x_0 - E[X] \right)^2 P(X = x_0) + \left(x_1 - E[X] \right)^2 P(X = x_1) \quad (73)$$

$$= (0 - p)^2 (1 - p) + (1 - p)^2 p \quad (74)$$

$$= p(1 - p)(p + 1 - p) \quad (75)$$

$$= \underline{\underline{p(1 - p)}} \quad (76)$$

altså variansen til $X \sim \text{Ber}[p]$ er $p(1 - p)$.

d) Siden vi kjenner variansen så er standardavviket lett å finne:

$$\underline{\underline{\sigma[X]}} = \sqrt{Var[X]} = \underline{\underline{\sqrt{p(1 - p)}}} \quad (77)$$

e) Forsøksserien som Karo Pharma utfører har følgende **egenskaper**:

1. Hvert forsøk (pasient) har kun 2 mulige utfall, medisinen bedre (suksess) eller ikke bedre (fiasko) enn paracet
2. I oppgaven antas det at det er **samme sannsynlighet** p ($= 0.80$) i alle n ($= 10$) forsøkene.
3. Alle forsøk er **uavhengige**.
4. Vi gjennomfører et bestemt antall forsøk, $n = 10$ i dette tilfellet.

Forsøksserien oppfyller dermed kravene til en **binomisk** forsøksserie.
Den stokastiske variabelen

$$Y = X_1 + X_2 + \dots X_{10} \tag{78}$$

beskriver antallet suksesser i denne forsøksserien er derfor **binomisk fordelt**.

- f) Forsøkene er uavhengige.
Sannsynligheten for 9 suksesser for en (hvilken som helst) gitt sekvens er gitt ved ⁴

$$P(\text{gitt mulig sekvens}) = p^9 (1 - p)^1 \quad (79)$$

Men 9 suksesser kan fås på flere forskjellige måter:

$$\underbrace{\text{sssssssssf}, \text{ssssssssfs}, \text{sssssssfss}, \dots, \text{fsssssssss}}_{= \text{totalt 10 mulige måter/sekvenser å få 9 suksesser på}} \quad (80)$$

Alle disse sekvensene har samme sannsynlighet.

Dermed kan vi bruke “**urnemodellen**” for å telle opp antall mulige måter å få 9 suksesser på (av totalt 10 forsøk).

Egenskapene til forsøksserien i lign.(80) er:

- siden sannsynligheten p er uendret etter trekking så tilsvarer det et valg *uten* tilbakelegging
- rekkefølge har ikke betydning, dvs. *ikke*-ordnet utvalg

Dette er derfor et *ikke*-ordnet utvalg *uten* tilbakelegging, dvs. **situasjon 3**.

Dermed gjelder:

$$\# \text{ ikke-ordnede komb. u/tilbakelegging} = \frac{10!}{(10-9)! 9!} \equiv \underbrace{\binom{10}{9}}_{\text{binomialkoeff.}} = 10 \quad (81)$$

hvor

$$n = 10 \quad (\text{antall valgobjekter/trekk}) \quad (82)$$

$$y = 9 \quad (\text{antall objekter som velges}) \quad (83)$$

⁴I lign.(79) bruker vi den spesielle multiplikasjonssetningen.

Siden forsøkene er uavhengige så gjelder den **spesielle addisjonssetningen**:

$$\underline{\underline{P(Y = 9)}} = P\left(\text{sssssssssf} \cup \text{ssssssssfs} \cup \dots \cup \text{fsssssssss}\right) \quad (84)$$

$$\stackrel{\text{spes. add.}}{=} P(\text{sssssssssf}) + P(\text{ssssssssfs}) + \dots + P(\text{fsssssssss}) \quad (85)$$

$$= p^9(1-p)^1 + p^9(1-p)^1 + \dots + p^9(1-p)^1 \quad (86)$$

$$= 10 p^9(1-p)^1 \quad (87)$$

$$= \begin{array}{l} \text{antall måter å få en gitt mulig sekvens på} \\ \times \\ P(\text{gitt mulig sekvens}) \end{array} \quad (88)$$

$$= \underline{\underline{\binom{10}{9} p^9 (1-p)^1}}, \quad \text{q.e.d.} \quad (89)$$

g) Kravet som Weifa stiller kan formuleres som:

$$P(Y = 10) = 0.80 \quad (90)$$

For $P(Y = 10)$ så bruker vi formelen som oppgitt i oppgaven med $n = 10$ og $y = 10$:

$$\binom{n}{y} p^y (1-p)^{n-y} = 0.80 \quad (91)$$

$$\binom{10}{10} p^{10} (1-p)^0 = 0.80 \quad (92)$$

$$p^{10} = 0.80 \quad (93)$$

$$p = 0.80^{1/10} \quad (94)$$

$$\underline{\underline{p = 0.9780}} \quad (95)$$

h) Tolkning:

$$\begin{aligned} \underline{\underline{p}} &= 0.9780 \\ &= \text{suksess-sannsynligheten for i et gitt Bernoulli-forsøk for at} \end{aligned} \quad (96)$$

det skal være 80 % sannsynlighet for at 10 av 10 pasienter opplever suksess

■