

LÄS FÖRST DETTA:

- **BESVARA UPPGIFT 1 SAMT TRE AV UPPGIFTERNA 2-5 (TOTALT FYRA UPPGIFTER)! Om du besvarat 5 frågor fälls den sista i ordningen bort från bedömningen.**
- **Besvara var och en av de fyra uppgifterna på en egen sida = 4 sidor = 1 ark. ANVÄND EXTRA ARK ENDAST SOM KLOTTPAPPER.**
- **Alla svar skall innehålla motiveringar, uträkningar eller uppställningar. Ange tydligt ditt slutliga svar genom att streckा under det.**
- **Varje uppgift ger max 10 poäng. För godkänd tent krävs minst 20 poäng.**
- **Om du skall delta i FUM-kursen skriv "FUM" synligt överst på ditt papper, så att ditt resultat GENAST kan registreras i Oodi.**

1. (Besvaras av alla)

I en population är såväl kvinnors som mäns längder normalfördelade. Kvinnorna är i genomsnitt 168 centimeter långa medan männen i genomsnitt är 178 cm. För både män och kvinnor är standardavvikelsen 6 cm.

- a) Hur stor är sannolikheten att en slumpmässigt vald kvinna är mellan 160 och 170 centimeter lång?
- b) En man och en kvinna väljs slumpmässigt, hur stor är sannolikheten att kvinnan är längre än mannen?
(Notera att för skillnader gäller att: $\mu_{x_1-x_2} = \mu_{x_1} - \mu_{x_2}$ och $\sigma^2_{x_1-x_2} = \sigma^2_{x_1} + \sigma^2_{x_2}$.)

2. Johan är ute och fiskar och han får i genomsnitt två abborrar per timme. Antalet fångade abborrar per tidsenhet kan approximeras med en Poissonfördelning.

- a) Hur stor är sannolikheten att Johan får minst två abborrar under en timme?
- b) Hur stor är sannolikheten att Johan får minst två abborrar under två timmar?
- c) Ange väntevärde och varians för antalet fångade abborrar under fyra timmar.

3. Ordningsreglerna för ett husbolag säger att man får ha högst en hund och en katt per hushåll. Av de 50 hushållen i bolaget har 25 en hund, 7 har både en katt och en hund och 18 har varken en katt eller en hund. Om man slumpmässigt väljer ut ett hushåll, hur stor är sannolikheten att hushållet har

- a) en katt?
- b) en hund men ingen katt?
- c) en katt eller en hund men inte båda?

4. Ett företag har produktionsenheter i två olika städer. I frekvenstabellen nedan ser du de heltidsanställdas lönefördelning.

| Månadslön | Frekvens, enhet Kemi | Frekvens, enhet Salo |
|----------------|----------------------|----------------------|
| Mindre än 2500 | 0 | 10 |
| 2500 – 2699,99 | 3 | 16 |
| 2700 – 2899,99 | 10 | 33 |
| 2900 – 3099,99 | 43 | 97 |
| 3100 – 3299,99 | 69 | 120 |
| 3300 – 3499,99 | 93 | 136 |
| 3500 – 3699,99 | 111 | 150 |
| 3700 – 3899,99 | 129 | 100 |
| 3900 – 4099,99 | 120 | 42 |
| 4100 – 4299,99 | 60 | 20 |

Besvara följande frågor, motivera kort dina svar.

- a)** Vilken enhet har högst medellönen? Hur hög är den?
- b)** Vilken enhet har högst medianlön? Hur hög är den?
- c)** Vilken enhet har högst typvärde för månadslönen? Hur hög är den?

5. En turist har gjort upp en lista över tio sevärdheter han vill besöka, men han märker att han bara har tid att besöka sex av dem. På hur många olika sätt kan detta göras,

- a)** om man beaktar i vilken ordning sevärdheterna besöks?
- b)** om man inte beaktar i vilken ordning sevärdheterna besöks?

Anta ytterligare att sex av sevärdheterna finns i stad A medan fyra finns i stad B och vår turist bestämmer sig för att se tre sevärdheter i vardera staden.

- c)** På hur många olika sätt kan detta göras om han först besöker sevärdheterna i stad A, sedan sevärdheterna i stad B? Ange svar både med och utan att beakta ordningsföljden.

Lägesmått för stickprov

- **Modus:** M_0 = typvärdet med den största frekvensen $f(x)$
- **Median:** M_d = värdet på den mittersta observationen i ett storleksordnat material. Om antalet observationer är jämnt, är M_d ett genomsnitt av de två mittersta värdena.
- **Medeldvärdet:** $\bar{x} = \frac{\sum_{i=1}^n x_i}{n} = \frac{x_1 + x_2 + \dots + x_n}{n}$
- **Klassindelat material:** \bar{x} är mitten i respektive klass vid beräkning av lägesmåttet.
- **Modus:** M_0 = klassmitten i klassen med största frekvens,
- **Median:** M_d = klassmitten i klassen där den "mittersta" observationen finns i det storleksordnade materialet, d.v.s. där $P(X)$ kommer upp till 50%
- **Medeldvärdet:** $\bar{x} = \frac{\sum_{k=1}^n f_k x_k}{\sum_{k=1}^n f_k}$ =summan av klassmittarna x_k vägda med respektive klassfrekvens f_k

- **Kvartilsavvikelse:** $Q = \frac{Q_3 - Q_1}{2}$ = ett halvt kvartilavstånd $Q_3 - Q_1 \Rightarrow$ medelavståndet mellan värdena på de observationer för vilka $P(X) = 25\%$ och $P(X) = 75\%$ i ett storleksordnat material.
- **Variationsvidd, variationsbredd (eng. range):** $R = x_{\max} - x_{\min}$
- **Standardavvikelse:** $s = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}}$. **Varians:** s^2
- **Klassindelat material:** används i regel klassmittarna i respektive klass vid beräkning av spridningsmåttet:

- **Kvarthälfte:** $Q =$ det genomsnittliga avståndet mellan klassmittarna i de klasser där Q_1 och Q_3 finns (se ovan).
- **Variationsvidd:** R anger i allmänhet som avståndet mellan de yttersta klassgränserna
- **Standardavvikelse:** $s = \sqrt{\frac{\sum_{k=1}^n (x_k - \bar{x})^2}{n-1}}$ = summan av klassmittarnas kvadrerade avstånd från medeldvärdet vägda med respektive klassfrekvens. **Varians:** s^2

Formler för 1112-2 EMS-statistik / 1

Diskreta sannolikhetsfördelningar (allmänt)
En diskret stokastisk variabel X antar värdena x_1, x_2, \dots, x_n med sannolikheterna p_1, p_2, \dots, p_n .

| Fördelning | Definitionsängd | Discret sannolikhet | Väntevärde | Varians |
|----------------------------|--|---|---------------|---|
| Binomial-fördelning | $X \sim \text{Bin}(n, \pi)$ $r = 0, 1, \dots, n$ | $P(X = r) = \binom{n}{r} \cdot \pi^r \cdot (1-\pi)^{n-r}$ | $n \cdot \pi$ | $n \cdot \pi \cdot (1-\pi) \frac{N-n}{N-1}$ |
| Hypergeometrisk fördelning | $X \sim \text{Hyp}(N, N_x, n)$ $r = 0, 1, \dots, n$ | $P(X = r) = \frac{\binom{N_x}{r} \cdot \binom{N-N_x}{n-r}}{\binom{N}{n}}$ | $n \cdot \pi$ | $n \cdot \pi \cdot (1-\pi) \frac{N-n}{N-1}$ |
| Poisson-fördelning | $X \sim \text{Po}(\mu)$ $r = 0, 1, 2, \dots$ | $P(X = r) = e^{-\mu} \cdot \frac{\mu^r}{r!}$ | \bar{x} | μ |

Kontinuerliga sannolikhetsfördelningar (allmänt)

Om en kontinuerlig stokastisk variabel X är definierad i utfallsrummet $\Omega = [a, b]$, där delintervall utgör händelser, t.ex. $c \leq X \leq d$ där $[c, d] \subseteq \Omega$, så gäller för dess tätförfunktion $f(x)$:

- $f(x) \geq 0$ för alla x
- **Totalsannolikhet** $\int_a^b f(x) dx = 1$ där $x \in [a, b]$ (om $\Omega = \mathbb{R}$, så är $\int_{-\infty}^{\infty} f(x) dx = 1$)
- **Fördelningsfunktion** $F(x) = P(X \leq x) = \int_a^x f(t) dt$
- **Sannolikhet för en händelse** $P(A) = P(c \leq X \leq d) = \int_c^d f(x) dx = F(d) - F(c)$
- **Väntevärde** $E(x) = \int_a^b f(x) \cdot x dx$ där $x \in [a, b]$
- **Varians** $V(x): \sigma^2 = \int_a^b f(x) \cdot (x - \mu)^2 dx$ där $x \in [a, b]$
- **Standardavvikelse** $\sigma = \sqrt{\sigma^2}$

| Kombinatorik | Formel | Påräknen |
|---|--|------------------------------------|
| Antal permutitioner (ordningsföljder) av n element | $n! = n \cdot (n-1) \cdot (n-2) \cdots 1$ | $n!$ |
| Antal permutitioner av r urval (urval med ordningsföljd) | $P_r^n = \frac{n!}{(n-r)!}$ | nPr |
| där r element väljs ur en mängd av n element | $C_r^n = \binom{n}{r} = \frac{n!}{r!(n-r)!}$ | nCr |
| Antal kombinationer av r urval/då r element väljs ur en mängd av n element (oberoende av ordning) | $C_r^{n,r} = \frac{(n+r-1)!}{r!(n-1)!}$ | nCr , där n ges värdet $n+r-1$ |
| Antal konfigurationer av n element med uppreppning | | |

Union: $P(A \text{ eller } B \text{ eller båda}) = P(A \cup B) = P(A) + P(B) - P(A \cap B)$
Snitt: $P(A \text{ och } B) = P(A \cap B) = P(A)P(B)$ om A och B är oberoende, annars: $P(A \cap B) = P(A)P(B|A)$.
Bayes satz: $P(A_1 | B) = \frac{P(A_1)P(B | A_1)}{P(B)}$ $= \frac{P(A_1)P(B | A_1) + P(A_2)P(B | A_2) + \dots + P(A_n)P(B | A_n)}{P(B)}$

Formler för 1112-2 EMS-statistik / 2

Diskreta sannolikhetsfördelningar (allmänt)
En diskret stokastisk variabel X antar värdena x_1, x_2, \dots, x_n med sannolikheterna p_1, p_2, \dots, p_n .

| Fördelning | Definitionsängd | Kontinuerlig sannolikhet (intervall) | Väntevärde | Varians |
|----------------------|---|---|-----------------|----------------------|
| Normal-fördelning | $X \sim N(\mu, \sigma^2)$ $X \in \mathbb{R}$ | Den kumulativa sannolikheten: $P(X \leq x_0) = P(Z \leq \frac{x_0 - \mu}{\sigma}) = \Phi(z_0)$ finns tabellförd på sista sidan. | μ | σ^2 |
| Likformig fördelning | $X \sim U(a, b)$ $X \in [a, b]$ | $P(c \leq X \leq d) = \frac{d-c}{b-a}$ | $\frac{a+b}{2}$ | $\frac{(b-a)^2}{12}$ |

4. Ett företag har produktionsenheter i två olika städer. I frekvenstabellen nedan ser du de heltidsanställdas lönefördelning.

| Månadslön | Frekvens, enhet Kemi | Frekvens, enhet Salo |
|----------------|----------------------|----------------------|
| Mindre än 2500 | 0 | 10 |
| 2500 – 2699,99 | 3 | 16 |
| 2700 – 2899,99 | 10 | 33 |
| 2900 – 3099,99 | 43 | 97 |
| 3100 – 3299,99 | 69 | 120 |
| 3300 – 3499,99 | 93 | 136 |
| 3500 – 3699,99 | 111 | 150 |
| 3700 – 3899,99 | 129 | 100 |
| 3900 – 4099,99 | 120 | 42 |
| 4100 – 4299,99 | 60 | 20 |

Besvara följande frågor, motivera kort dina svar.

- a)** Vilken enhet har högst medellönen? Hur hög är den?
- b)** Vilken enhet har högst medianlön? Hur hög är den?
- c)** Vilken enhet har högst typvärde för månadslönen? Hur hög är den?

5. En turist har gjort upp en lista över tio sevärdheter han vill besöka, men han märker att han bara har tid att besöka sex av dem. På hur många olika sätt kan detta göras,

- a)** om man beaktar i vilken ordning sevärdheterna besöks?
- b)** om man inte beaktar i vilken ordning sevärdheterna besöks?

Anta ytterligare att sex av sevärdheterna finns i stad A medan fyra finns i stad B och vår turist bestämmer sig för att se tre sevärdheter i vardera staden.

- c)** På hur många olika sätt kan detta göras om han först besöker sevärdheterna i stad A, sedan sevärdheterna i stad B? Ange svar både med och utan att beakta ordningsföljden.

Lägesmått för stickprov

Formler för 1112-2 EMS-statistik / 1

- Modus:** $Mo = \text{typröde} = \text{värdet med den största frekvensen } f(x)$
- Median:** $Md = \text{värdet på den mittersta observationen i ett storleksordnat material}$. Om antalet observationer är jämt, är Md ett genomsnitt av de två mittersta värdena.

$$\text{Medelvärdet: } \bar{x} = \frac{\sum_{i=1}^n x_i}{n} = \frac{x_1 + x_2 + \dots + x_n}{n}$$

I klassindelat material används i regel klassmitttama i respektive klass vid beräkning av lägesmåttet:

- Modus:** $Mo = \text{klassmitten i klassen med största frekvens, storleksordnade materialet, d.v.s. där } P(x) \text{ kommer upp till } 50\%$
- Median:** $Md = \text{klassmitten i klassen där den "mittersta" observationen finns i det storleksordnade materialet, t.ex. s.d. } P(x) \text{ kommer upp till } 50\%$

$$\text{Medelvärdet: } \bar{x} = \frac{\sum_{k=1}^n f_k x_k}{n} = \text{summan av klassmittarna } x_k \text{ vägda med respektive klassfrekvens } f_k$$

Spridningsmått för stickprov

- Kvarтиlivarieté:** $Q = \frac{Q_3 - Q_1}{2} = \text{ett halvt kvarтиlivstånd } Q_3 - Q_1 \Rightarrow \text{medelavståndet mellan värdena på de observationer för vilka } P(x)=25\% \text{ och } P(x)=75\% \text{ i ett storleksordnat material.}$
- Variationsvidd, variationsbredd (eng. range):** $R = x_{\max} - x_{\min}$

$$\text{Standardavvikelse: } s = \sqrt{\frac{\sum_{k=1}^n f_k (x_k - \bar{x})^2}{n-1}}. \quad \text{Varians: } s^2$$

I klassindelat material används i regel klassmitttama i respektive klass vid beräkning av spridningsmåttet:

- Kvarтиlivarieté:** $Q = \text{det genomsnittliga avståndet mellan klassmittarna i de klasser där } Q_1 \text{ och } Q_3 \text{ finns (se ovan).}$
- Variationsvidd:** R anger allmänhet som avståndet mellan de yttersta klassgränserna medelvärdet vägda med respektive klassfrekvens. $\text{Varians: } s^2$

$$\text{Standardavvikelse: } s = \sqrt{\frac{\sum_{k=1}^n f_k (x_k - \bar{x})^2}{n-1}} = \text{summan av klassmittarnas kvadrerade avstånd från medelvärdet vägda med respektive klassfrekvens. Varians: } s^2$$

| Kombinatorik | Antal permutации (ordningsföljder) av n element | Formel | Påräknaren | Kontinuerlig sannolikhetsfördelning | Kontinuerlig sannolikhetsfördelning (interval) | Väntevärde | Varians |
|--|---|---------------------------------------|------------|---|--|------------|------------|
| Antal permutации av <u>urval</u> (urval med ordningsförfat) | $n! = n \cdot (n-1) \cdot (n-2) \cdots 1$ | $n!$ | nPr | $X \sim N(\mu, \sigma^2)$ $X \in \mathbb{R}$ | $\int_a^b f(x) \cdot x \, dx$ där $x \in [a, b]$ | μ | σ^2 |
| då r element väljs ur en mängd av n element | $P_r^n = \frac{n!}{(n-r)!}$ | | | $V(X): \sigma^2 = \int_a^b f(x) \cdot (x - \mu)^2 \, dx$ där $x \in [a, b]$ | $P(Z \leq \frac{X_0 - \mu}{\sigma}) = \Phi(z_0)$ | | |
| Antal kombinationer av <u>urval</u> då r element väljs ur en mängd av n element (oberoende av ordning) | $C_r^n = \binom{n}{r} = \frac{n!}{r!(n-r)!}$ | nCr | | $\text{Standardavvikelse: } \sigma = \sqrt{\sigma^2}$ | | | |
| Antal kombinationer av <u>urval</u> då r element väljs ur en mängd av n element med upprepning | $C'_r^n = \frac{(n+r-1)!}{r!(n-1)!}$ | nCr , där n ges värdet $n+r-1$ | | | | | |

$$\text{Union: } P(A \text{ eller } B \text{ eller båda}) = P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

$$\text{Schnitt: } P(A \text{ och } B) = P(A \cap B) = P(A)P(B)$$

$$\text{Bayes sats: } P(A_1 | B) = \frac{P(A_1 \text{ och } B)}{P(B)} = \frac{P(A_1)P(B | A_1)}{P(A_1)P(B | A_1) + P(A_2)P(B | A_2) + \dots + P(A_n)P(B | A_n)}$$

Diskreta sannolikhetsfördelningar (allmänt)

En diskret sannolikhet är variabel X antar värden x_1, x_2, \dots, x_n med sannolikheterna p_1, p_2, \dots, p_n .

| Fördelning | Definitionsängd | Diskret sannolikhet | Väntevärde | Varians |
|----------------------------|--|---|--|---|
| Binomial-fördelning | $X \sim \text{Bin}(n, \pi)$ $r = 0, 1, \dots, n$ | $P(X = r) = \binom{n}{r} \cdot \pi^r \cdot (1-\pi)^{n-r}$ | $n \cdot \pi$ | $n \cdot \pi \cdot (1-\pi)$ |
| Hypergeometrisk fördelning | $X \sim \text{Hyp}(N, N_x, n)$ $r = 0, 1, \dots, n$ | $P(X = r) = \frac{\binom{N_x}{r} \cdot \binom{N-N_x}{n-r}}{\binom{N}{n}}$ | $\mu \approx n \cdot \pi$ där $\pi = \frac{N_x}{N}$ | $n \cdot \pi \cdot (1-\pi) \cdot \frac{N-n}{N-1}$ |
| Poisson-fördelning | $X \sim \text{Po}(\mu)$ $r = 0, 1, 2, \dots$ | $P(X = r) = e^{-\mu} \cdot \frac{\mu^r}{r!}$ | μ | μ |

Kontinuerliga sannolikhetsfördelningar (allmänt)

Om en kontinuerlig stokastisk variabel X är definierad i utfallsummet $\Omega = [a, b]$, där delintervall utgör händelser, t.ex. $c < X \leq d$ där $[c, d] \subseteq \Omega$, så gäller för dess tätfördelningsfunktion $f(x)$:

- $f(x) \geq 0$ för alla x
- $\text{Totalsannolikhet: } \int_a^b f(x) dx = 1$ då $x \in [a, b]$ (om $\Omega = \mathbb{R}$, så är $\int_a^b f(x) dx = 1$)
- $\text{Fördelningsfunktion: } F(x) = P(X \leq x) = \int_a^x f(t) dt$
- $\text{Sannolikhet för en händelse: } P(A) = P(c \leq X \leq d) = \int_c^d f(x) dx = F(d) - F(c)$

| Kombinatorik | Formel | Påräknaren | Kontinuerlig sannolikhetsfördelning | Kontinuerlig sannolikhetsfördelning (interval) | Väntevärde | Varians |
|--|--|---------------------------------------|--|--|------------|------------|
| Antal permutации (ordningsföljder) av n element | $n! = n \cdot (n-1) \cdot (n-2) \cdots 1$ | $n!$ | $\int_a^b f(x) \cdot x \, dx$ där $x \in [a, b]$ | $P(X \leq x_0) = P(Z \leq \frac{x_0 - \mu}{\sigma}) = \Phi(z_0)$ | μ | σ^2 |
| Antal permutации av <u>urval</u> (urval med ordningsförfat) | $P_r^n = \frac{n!}{(n-r)!}$ | nPr | | | | |
| då r element väljs ur en mängd av n element | | | | | | |
| Antal kombinationer av <u>urval</u> då r element väljs ur en mängd av n element (oberoende av ordning) | $C_r^n = \binom{n}{r} = \frac{n!}{r!(n-r)!}$ | nCr | | | | |
| Antal kombinationer av <u>urval</u> då r element väljs ur en mängd av n element med upprepning | $C'_r^n = \frac{(n+r-1)!}{r!(n-1)!}$ | nCr , där n ges värdet $n+r-1$ | | | | |

Normalfördelningstabell (negativa värden)

Normalfördelningstabell (positiva värden)

| Z | $P(Z < z)$ |
|------|------------|
| -3.4 | 0.000394 |
| -3.3 | 0.000448 |
| -3.2 | 0.000659 |
| -3.1 | 0.00097 |
| -3.0 | 0.00135 |
| -2.9 | 0.00187 |
| -2.8 | 0.00256 |
| -2.7 | 0.00347 |
| -2.6 | 0.00466 |
| -2.5 | 0.00621 |
| -2.4 | 0.00820 |
| -2.3 | 0.01072 |
| -2.2 | 0.01350 |
| -2.1 | 0.01766 |
| -2.0 | 0.02275 |
| -1.9 | 0.02878 |
| -1.8 | 0.03553 |
| -1.7 | 0.04456 |
| -1.6 | 0.05480 |
| -1.5 | 0.06681 |
| -1.4 | 0.08076 |
| -1.3 | 0.09680 |
| -1.2 | 0.11507 |
| -1.1 | 0.13666 |
| -1.0 | 0.15865 |
| -0.9 | 0.18406 |
| -0.8 | 0.21195 |
| -0.7 | 0.24196 |
| -0.6 | 0.27425 |
| -0.5 | 0.30853 |
| -0.4 | 0.34457 |
| -0.3 | 0.38209 |
| -0.2 | 0.42074 |
| -0.1 | 0.46017 |
| 0.0 | 0.50000 |
| 0.1 | 0.53983 |
| 0.2 | 0.57926 |
| 0.3 | 0.61791 |
| 0.4 | 0.65542 |
| 0.5 | 0.69448 |
| 0.6 | 0.73275 |
| 0.7 | 0.76944 |
| 0.8 | 0.80459 |
| 0.9 | 0.83859 |
| 1.0 | 0.86642 |
| 1.1 | 0.89234 |
| 1.2 | 0.91624 |
| 1.3 | 0.93793 |
| 1.4 | 0.95759 |
| 1.5 | 0.97537 |
| 1.6 | 0.99149 |
| 1.7 | 0.99595 |
| 1.8 | 0.99834 |
| 1.9 | 0.99944 |
| 2.0 | 0.99971 |
| 2.1 | 0.99981 |
| 2.2 | 0.99986 |
| 2.3 | 0.99990 |
| 2.4 | 0.99993 |
| 2.5 | 0.99995 |
| 2.6 | 0.99996 |
| 2.7 | 0.99997 |
| 2.8 | 0.99998 |
| 2.9 | 0.99999 |
| 3.0 | 0.99999 |
| 3.1 | 0.99999 |
| 3.2 | 0.99999 |
| 3.3 | 0.99999 |
| 3.4 | 0.99999 |

Källa: Wahlin, K: Tillämpad statistik – en grundkurs (2011)

Normalfördelningstabellen visar det z -värde som svarar mot en given area. Figuren illustrerar att 3.92 procent av fördelningen ligger till vänster om punkten $z = -1.76$.

Normalfördelningstabellen visar det z -värde som svarar mot en given area. Figuren illustrerar att 97.5 procent av fördelningen ligger till vänster om punkten $z = 1.96$.

0.975

0.9973

0

1.96

Formler för 1112-2 EMS-statistik / 3

0.9973

0

1.96