eksamen_2021v

Grunnleggende

a) Bernoulli-prosess

La A være alle R-brukere i verden, og B være alle MatLab- brukere. Bruk følgende anslag på antallene: R: |A| = 2~000~000~MatLab: |B| = 4~000~000~Overlapp: |AB| = 400~000

i.

La p1 være andelen R-brukere som også er MatLab-brukere, og la p2 være andelen MatLab-brukere som også er R-brukere. Regn ut disse to tallene.

ii.

På den nye jobben din er det 15 R-brukere. La X være hvor mange av dem som også er MatLab-brukere, og la X $\sim f(x)$. Finn a. f(x) b. E

X

c. Tegn sannsynlighetsfordelingen (pdf).

iii.

Du er på konferanse for MatLab-brukere. La Y være antallet ikke-R-brukere du hilser på før du har hilst på 4 R-brukere, og la Y $\sim g(y)$. Finn A. g(y) B. P(Y > 20) C. $\sigma^2 Y$.

b) Gaussisk prosess

i.

 $X \sim f(x)$ er en kontinuerlig stokastisk variabel med E[X] = 4, $\sigma X = 2$. Bruk normaltilnærmingen til f(x) til å regne ut $P(X \le 3)$.

$$X = \Phi_{(4,2)}(3) = 0.30853$$

ii.

 $Y \sim g(y)$ er en diskret stokastisk variabel med E[Y] = 4, $\sigma Y = 2$. Bruk normaltilnærmingen til g(y) til å regne ut $P(Y \le 3)$.

$$Y = \Phi_{(4,2)}(3 + \frac{1}{2}) = 0.4012937$$

Inferens

Gaussisk prosess

```
## [1] -1.387893 -0.927109 -4.962876 2.135971 -0.284439 13.543366 1.608963
## [8] -8.369690 -2.827972
## [1] 0.1587112
```

i.

$$\tau \sim \gamma_{(\frac{8}{2},\frac{295.879206}{2})}(t) = \tau \sim \gamma_{(4,147.939603)}(t)$$

$$\mu = t_{(-0.1635199, 2.0271738, 8)}(x)$$

$$X_{+} = t_{(-0.1635199, 6.4104863, 8)}(x)$$

ii.

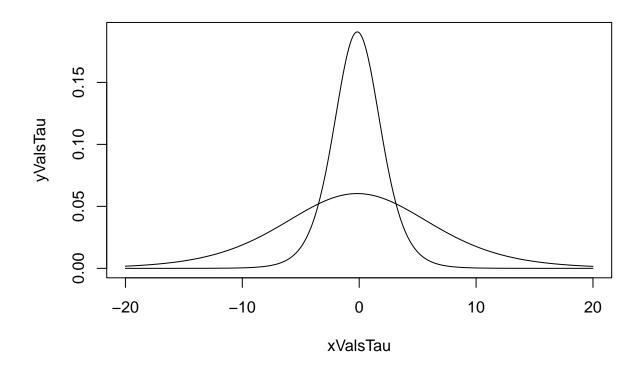
$$P(\sigma \leq 5) = P(\sigma^2 \leq 5^2) = P(\frac{1}{\sigma^2} \geq \frac{1}{5^2}) = P(\tau \geq 0.04)$$

$$P(\tau \ge 0.04) = 1 - \Gamma_{(4,147.939603)}(0.04) = 0.1587112$$

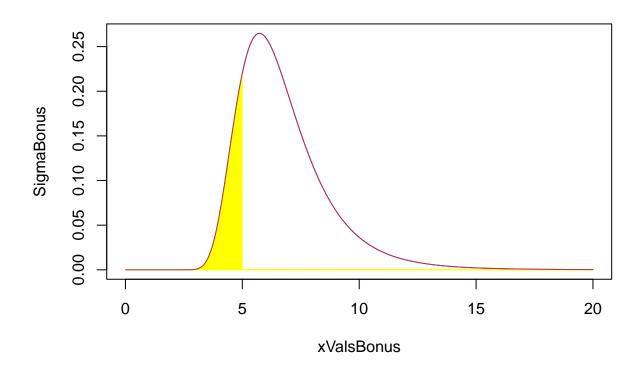
iii.

 H_0 blir da $P(\sigma \le 5)$, og H_1 blir da $P(\sigma > 5)$ Hvis α er større enn 0.1587112 så forkaster vi H_0 til fordel for H_1

vi.

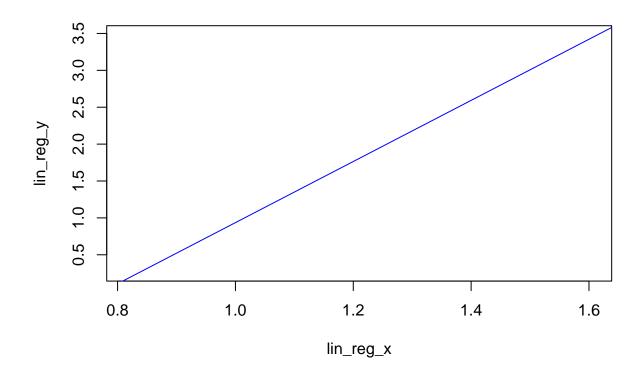


 $\mathbf{v}.$

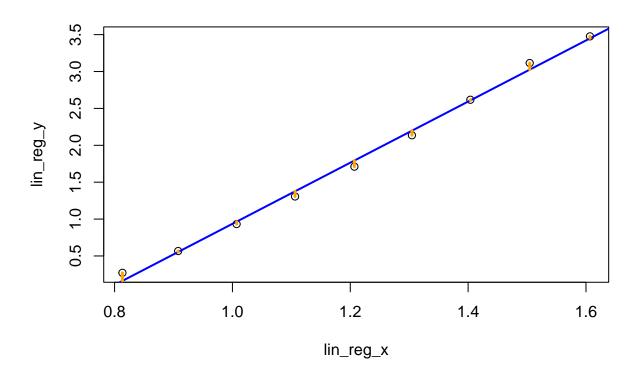


c) Lineær regresjon

(Intercept) lin_reg_x ## -3.206885 4.142698



i.



###iii.