

**UNIVERSITETET I OSLO**  
**ØKONOMISK INSTITUTT**

Eksamen i: **ECON2130 – Statistikk 1**

Eksamensdag: 30.05.2018

**Sensur kunngjøres: 19.06.2018**

Tid for eksamen: kl. 14:30 – 17:30

Oppgavesettet er på 5 sider

Tillatte hjelpemidler:

- Alle trykte og skrevne hjelpemidler er tillatt.
- To kalkulatorer av typen: Aurora HC106, og/eller Casio FX-85EX.

Eksamen blir vurdert etter ECTS-skalaen. A-F, der A er beste karakter og E er dårligste ståkarakter. F er ikke bestått.



# Eksamen i ECON2130, vår 2018

## Instruks

Denne eksamenen består av to hovedoppgaver som hver tillegges lik vekt. Hvert delspørsmål tillegges lik vekt for hvert hovedspørsmål.

To tabeller som kan være nyttige er vedlagt bakerst i oppgaveteksten.

## I. Livet som oppfinner

- 1) En oppfinner jobber med å finne opp duppeditter. Vi skal her anta at antall nye duppeditter hun finner opp i løpet av et år kan beskrives ved en Poisson-fordeling med rate  $\lambda = 6$ . Diskuter om det er en fornuftig beskrivelse av en oppfinner.
- 2) Finn sannsynligheten for
  - a) At hun gjør mer enn 10 oppfinnelser i løpet av et år
  - b) At hun gjør mellom 7 og 9 oppfinnelser i løpet av et år
  - c) Ikke finner opp noe en måned
  - d) At hun finner opp noe i løpet av en måned, gitt at hun ikke fant opp noe i forrige måned

For de resterende oppgaver vil ikke egenskaper ved Poisson-fordelinger utover forventningen og variansen være nødvendige å kjenne til.

- 3) Oppfinner-livet gir lite sikker inntekt, så vår oppfinner har slått seg sammen med to andre kolleger i et oppfinner-kollektiv. Vår oppfinner finner som kjent opp duppeditter med en årlig rate  $\lambda_1 = 6$ , mens de to kollegene finner opp duppeditter med årlige rater  $\lambda_2 = 4$  og  $\lambda_3 = 7$ . Finn forventningen og variansen til årlig oppfinnelses-produksjon fra kollektivet.
- 4) I kollektivet deles alle inntektene likt på de tre medlemmene. En typisk oppfinnelse er verdt kr 100 000. Definer en stokastisk variabel  $Y_A$  som er oppfinnerens inntekt som selvstendig og en stokastisk variabel  $Y_K$  som er inntekten som kollektiv-medlem. Finn forventet inntekt med de to ordningene samt variansene til inntekten. Hva er de lønnsmessige fordelene og ulempene for oppfinneren av å være en del av kollektivet?
- 5) Forklar hvordan du vil lage et R-skript som simulerer andelen av oppfinnelsene fra kollektivet som er produsert av vår oppfinner i løpet av et år. Hvordan kan du bruke dette til å finne sannsynlighet for at hun i et år har stått for mindre enn  $1/3$  av oppfinnelsene i kollektivet?  
*Du kan enten skrive dette ut som R-kode eller pseudo-kode, men vær presis på hvilke trinn du vil gå gjennom*
- 6) Forklar hvordan du vil lage et R-skript som beregner sannsynligheten for at vår oppfinner har funnet opp 3 eller færre duppeditter i løpet av et år, gitt at oppfinnerkollektivet totalt fant opp 10 duppeditter.  
*Du kan enten skrive dette ut som R-kode eller pseudo-kode, men vær presis på hvilke trinn du vil gå gjennom*
- 7) Hvis kollektivet har mange medlemmer som alle har  $\lambda = 6$ , hvordan vil fordelingen til inntekten til vår oppfinner være? Som før antar vi at inntektene deles likt på alle medlemmer. Hva er

sannsynligheten for at hun tjener mindre enn kr 550 000 et år hvis kollektivet har 100 medlemmer?  
Du kan her se bort fra medlemmene med  $\lambda \neq 6$ .

- 8) De siste tre årene har oppfinneren vår produsert henholdsvis 3, 5 og 4 nye oppfinnelser. Hun er bekymret for at hun ikke er like produktiv som før. Gjennomfør en hypotesetest hvor du tester om hun er mindre produktiv på 5% signifikansnivå. Vær klar på hvilket hypotesepar (nullhypotese og alternativ hypotese) du tester og hvilke forutsetninger du gjør.

Følgende kan være nyttig, men vær klar på hvorfor det kan brukes:

```
n<-1e5
x1<-rpois(n,lambda=6)
x2<-rpois(n,lambda=6)
x3<-rpois(n,lambda=6)
q<-x1<=3&x2<=5&x3<=4
mean(q)

[1] 0.01926
```

## II. Fordommer

For å studere fordommer mot svarte delte en gruppe forskere en utvalg hvite amerikanske respondenter tilfeldig inn i en behandlings- og en kontrollgruppe. De som var i kontrollgruppa fikk lest opp følgende tekst:

*Jeg skal nå lese opp tre ting som noen ganger gjør folk irriterte og sinte. Etter at jeg har lest opp alle tre ønsker jeg at du forteller meg HVOR MANGE av dem som irriterte deg. Jeg vil ikke vite hvilke, bare HVOR MANGE.*

- *Myndighetene øker bensinavgiftene*
- *Idrettsutøvere får million-lønninger*
- *Store bedrifter forurenses miljøet*

Respondentene i behandlingsgruppa fikk lest opp samme tekst, men fikk i tillegg et ekstra punkt på sin liste:

- *En svart familie flytter inn i nabolaget*

Her er litt deskriptiv statistikk fra en data frame 'race', hvor variabelen 'y' er antall kulepunkter som ble oppgitt å være irriterende, 'treat' er en binær variabel som er 1 for de i behandlingsgruppa og 0 i kontrollgruppa og 'south' en tilsvarende indikatorvariabel for å bo i en sørstat.

```

> dstats = function(x) c(mean=mean(x), sd=sd(x), n=length(x))

> stats(race$y)
      mean      sd      n
2.15587349 0.86840818 1328.00000000

> aggregate(y ~ treat, race, dstats)
  treat  y.mean  y.sd  y.n
1 0      2.11901082 0.80424201 647.00000000
2 1      2.19089574 0.92448997 681.00000000

> aggregate(y ~ treat*south, race, dstats)
  treat south  y.mean  y.sd  y.n
1 0      0      2.17391304 0.80376617 483.00000000
2 1      0      2.18714556 0.90711893 529.00000000
3 0      1      1.95731707 0.78600184 164.00000000
4 1      1      2.20394737 0.98558361 152.00000000

```

- 1) Vil du forvente at noen *uten* rasefordommer svarer annerledes på spørsmålene i behandlingsgruppa enn spørsmålene i kontrollgruppa? Hva med en som har rasefordommer? Forklar svaret ditt.
- 2) Beregn andel av hele populasjonen som ser ut til å ha rasefordommer. Beregn standardfeilen og konstruer et 95 % konfidensintervall for estimatet. Gi en kort tolkning av resultatet.
- 3) Gjennomfør en tosidig hypotesetest hvor nullhypotesen er at andelen med rasefordommer i populasjonen er null. Beregn z-verdien og p-verdien. Gjennomfør deretter hypotesetesten med 5 % konfidensnivå. Hvilke forutsetninger må du gjøre for at testen skal være gyldig?
- 4) Forsett med nullhypotesen om ingen rasefordommer og 5 % konfidensnivå, men anta nå at den sanne andelen som har fordommer er 0.1. Hva er teststyrken med den utvalgsstørrelsen vi har? Anta at variansen i 'y' er 1 i den behandla gruppa og 0.8 i kontrollgruppa.
- 5) Beregn differansen mellom populasjonsandelen som utviser fordommer i sørstatene og i ikke-sørstatene, og konstruer et 95 % konfidensintervall for denne differansen. Tolk differansen og diskuter om den er signifikant. Skill mellom praktisk og statistisk signifikans.
- 6) Kan vi tolke forskjellen i andelen respondenter som viser rasefordommer mellom sørstatene og ikke-sørstatene kausalt – er den en effekt av å bo i en sørstat? Bruk kontrafaktiske scenarier i forklaringen din.



## Standard normal distribution

Values of the cumulative distribution function,  $\Pr(Z \leq z)$  where  $Z \sim N(0, 1)$ . For example,

$$\Pr(Z \leq 0.12) = \Pr(Z \leq 0.1 + 0.02) = 0.5478.$$

	0	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998
3.5	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998
3.6	0.9998	0.9998	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.7	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.8	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.9	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

### Tables of the Poisson Cumulative Distribution

The table below gives the probability of that a Poisson random variable  $X$  with mean  $= \lambda$  is less than or equal to  $x$ . That is, the table gives

$$P(X \leq x) = \sum_{r=0}^x \lambda^r \frac{e^{-\lambda}}{r!}$$

$\lambda =$	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0	1.2	1.4	1.6	1.8
$x = 0$	0.9048	0.8187	0.7408	0.6703	0.6065	0.5488	0.4966	0.4493	0.4066	0.3679	0.3012	0.2466	0.2019	0.1653
1	0.9953	0.9825	0.9631	0.9384	0.9098	0.8781	0.8442	0.8088	0.7725	0.7358	0.6626	0.5918	0.5249	0.4628
2	0.9998	0.9989	0.9964	0.9921	0.9856	0.9769	0.9659	0.9526	0.9371	0.9197	0.8795	0.8335	0.7834	0.7306
3	1.0000	0.9999	0.9997	0.9992	0.9982	0.9966	0.9942	0.9909	0.9865	0.9810	0.9662	0.9463	0.9212	0.8913
4	1.0000	1.0000	1.0000	0.9999	0.9998	0.9996	0.9992	0.9986	0.9977	0.9963	0.9923	0.9857	0.9763	0.9636
5	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9998	0.9997	0.9994	0.9985	0.9968	0.9940	0.9896
6	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9997	0.9994	0.9987	0.9974
7	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9997	0.9994
8	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999
9	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

  

$\lambda =$	2.0	2.2	2.4	2.6	2.8	3.0	3.2	3.4	3.6	3.8	4.0	4.5	5.0	5.5
$x = 0$	0.1353	0.1108	0.0907	0.0743	0.0608	0.0498	0.0408	0.0334	0.0273	0.0224	0.0183	0.0141	0.0097	0.0041
1	0.4060	0.3546	0.3084	0.2674	0.2311	0.1991	0.1712	0.1468	0.1257	0.1074	0.0916	0.0611	0.0404	0.0266
2	0.6767	0.6227	0.5697	0.5184	0.4695	0.4232	0.3799	0.3397	0.3027	0.2689	0.2381	0.1736	0.1247	0.0884
3	0.8571	0.8194	0.7787	0.7360	0.6919	0.6472	0.6025	0.5584	0.5152	0.4735	0.4335	0.3423	0.2650	0.2017
4	0.9473	0.9275	0.9041	0.8774	0.8477	0.8153	0.7806	0.7442	0.7064	0.6678	0.6288	0.5321	0.4405	0.3575
5	0.9834	0.9751	0.9643	0.9510	0.9349	0.9161	0.8946	0.8705	0.8441	0.8156	0.7851	0.7029	0.6160	0.5289
6	0.9955	0.9925	0.9884	0.9828	0.9756	0.9665	0.9554	0.9421	0.9267	0.9091	0.8893	0.8311	0.7622	0.6860
7	0.9989	0.9980	0.9967	0.9947	0.9919	0.9881	0.9832	0.9769	0.9692	0.9599	0.9489	0.9134	0.8666	0.8095
8	0.9998	0.9995	0.9991	0.9985	0.9976	0.9962	0.9943	0.9917	0.9883	0.9840	0.9786	0.9597	0.9319	0.8944
9	1.0000	0.9999	0.9998	0.9996	0.9993	0.9989	0.9982	0.9973	0.9960	0.9942	0.9919	0.9829	0.9682	0.9462
10	1.0000	1.0000	1.0000	0.9999	0.9998	0.9997	0.9995	0.9992	0.9987	0.9981	0.9972	0.9933	0.9863	0.9747
11	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9999	0.9998	0.9996	0.9994	0.9991	0.9976	0.9945	0.9890
12	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9999	0.9998	0.9997	0.9992	0.9980	0.9955
13	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9997	0.9993	0.9983
14	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9998	0.9994
15	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9998
16	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999
17	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000