Oppgåve 1

Ein berømt kornprodusent hevdar at det gjennomsnittlege sukkerinnhaldet i ei frukostblandingeske som han sel, er 0.2 gram per 1 gram frukostblanding. Eit tilfeldig utval av n frukostblandingesker blir valde ut og sukkerinnhaldet blir registrert. Vi kan generelt gå ut ifrå at alle målingar, $X_1, X_2, \ldots X_n$ er uavhengige og normalfordelte med forventingsverdi μ, som representerar sukkerinnhaldet i korna, og standardavvik $\sigma = 0.03$.

Ein vil utføre ein hypotesetest med nullhypotese $H_0 : \mu = 0.2$ mot den alternative hypotesen $H_1 : \mu > 0.2$, og forkaste nullhypotesen på eit signifikansnivå gitt ved 0.05.

a) Føreslå korleis ein slik hypotesetest basert på gjennomsnittet $\bar{X}$ av målingane kan gjennomførast. Kva blir konklusjonen dersom det blir teke $n = 10$ prøver og gjennomsnittet av målingane er 0.22?

b) Finn sannsynet for at nullhypotesen blir forkasta dersom det blir teke $n = 10$ prøver og sukkerinnhaldet er $\mu = 0.22$.

c) Gå ut i frå at sukkerinnholdet er $\mu = 0.22$. Rekn ut kor mange prøver, n, som må takast for at sannsynet for at nullhypotesen blir forkasta skal vere større enn 0.8.

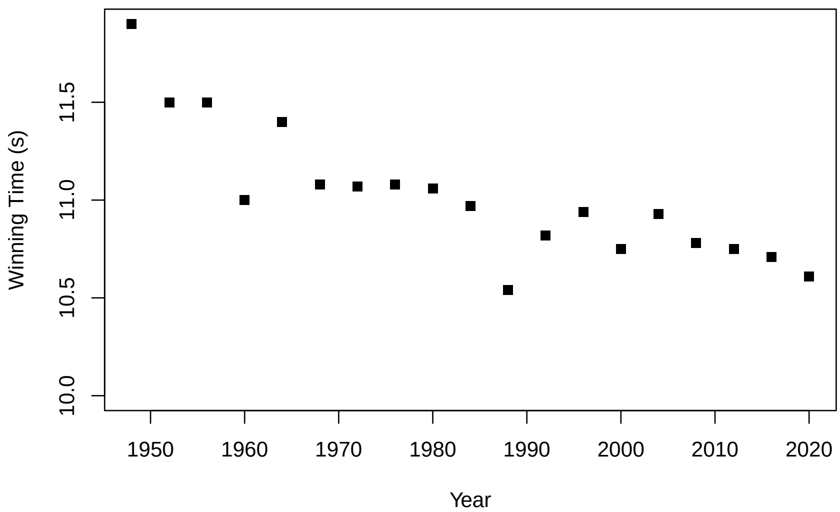
d) Eit nytt spørsmål oppstår om fargen på frukostblandinga påverkar salet. For å teste dette blei frukostblandinga selt i fire forskjellegfarga boksar. Talet på selde boksar for kvar farge den første veka var:

Blå = 31, Gul = 29, Rød = 23, Grøn = 29.

Bruk eit 5% signifikansnivå og test nullhypotesen om at talet på selde bokser av kvar av desse fire fargane er det same.

Oppgåve 2

Følgjande figur viser olympiske vinnartider i sekund på 100-meter for kvinner for dei olympiske åra frå 1948 til 2020:



Ein linear regresjonsmodell, $y_i = \beta_0 + \beta_1 x_i + \epsilon_i$ med $\epsilon_i \sim \mathcal{N}(0, \sigma^2)$, blei tilpassa vinnartidene. Her er y_i vinnartida i sekund og x_i gjev det olympiske året. Vi har totalt 19 observasjonspar slik at $i = 1, \ldots, 19$. Utrekning gav følgjande parameterestimat:

$$\hat{\beta}_0 = 36.78, \quad \hat{\beta}_1 = -0.013, \quad \hat{\sigma}^2 = 0.036,$$

Vidare er, $\sum_{i=1}^{n}(x_i - \bar{x})^2 = 9120$ (det vil sei 0.036/9120 er eit estimat av variansen til estimatoren til β_1).

a) Kva er endringa i forventa vinnartid frå den eine olympiaden til den neste (olympiske leikar blir arrangert kvart 4. år)? Kva ville ifølgje modellen vore vinnertida i år 0?

b) Kan vi vere sikre på at vinnartida blir raskare over tid? Føreslå og utfør ein hypotesetest for å løyse dette spørsmålet. Når du gjer dette, definerer du nullhypotesen og den alternative hypotesen, reknar ut teststatistikken og bestemmer det kritiske området for eit signifikansnivå på 5%. Kva konkluderar du med?

Oppgåve 3

Ein studie undersøkte om hundar føretrekker klapping eller vokal ros. Forskarar plasserte 14 hundar tilfeldig i to grupper med 7 i kvar. I gruppe A ville eigaren klappe hunden, medan i gruppe B ville eigaren gje lokal ros. Responsvariabelen er tida, i sekund, som hunden samhandla med eigaren sin og er gitt som:

Gruppe A: X_i	114	203	217	254	256	284	296
Gruppe B: Y_j	4	7	24	25	48	71	294

med $i = 1, \ldots, n_A$ og $j = 1, \ldots, n_B$, med $n_A = n_B = 7$.

Vi testar

$$H_0 : f_X(x) = f_Y(x) \quad \text{for all } x \in (-\infty, +\infty)$$
$$H_1 : f_X(x) = f_Y(x - c) \quad \text{for all } x \in (-\infty, +\infty)$$

for ein vilkårleg konstant $c \neq 0$.

a) Rekn ut teststatistikken og p-verdien for ein Wilcoxons to-utvalstest (The Wilcoxon Rank Sum Test). Kva blir konklusjon din?

Oppgåve 4

Det siste året har det vore av interesse å følgje med i korleis inflasjonen har utvikla seg og i kva kategoriar den har vore størst. I ei undersøking innanfor daglegvarehandelen blei det spesielt sett på bakevarer, fersk frukt og grønnsaker, hermetisk frukt og grønnsaker, kjøtt og kategorien meieriprodukt. Innanfor kvar av desse 5 kategoriane blei 10 produkt plukka ut tilfeldig og prosentvis prisauke i ein periode på 2 månader for kvart produkt blei registrert. Du kan i resten av oppgåva gå ut i frå følgjande modell for den observerte prosentvise prisauken:

$$Y_{ij} = \mu_j + \epsilon_{ij}, \quad i = 1, 2, \ldots, 10; \quad j = 1, 2, \ldots, 5,$$

der $\epsilon_{ij} \sim \mathcal{N}(0, \sigma^2)$ og uavhengige.

Frå dei observerte dataane fikk ein følgjande verdiar for total kvadratsum og kvadratsummen for kategoriar:

$$\text{SSTOT} = \sum_{j=1}^{5} \sum_{i=1}^{10} (y_{ij} - \bar{y}_{..})^2 = 27.5$$

$$\text{SSTR} = \sum_{j=1}^{5} 10(\bar{y}_{.j} - \bar{y}_{..})^2 = 4.6$$

a) Konstruer den komplette variansanalysetabellen og føreta ein test på nullhypotesa om at forventa prisstigning er den same innanfor kvar kategori. Kva blir konklusjonen med eit signifikansnivå valgt til 0.05?

Gjennomsnittleg observert auke for kvar kategori blei: 0.1% for bakevarer, 0.5% for fersk frukt og grønnsaker, 0.6% for hermetisk frukt og grønnsaker, 1% for kjøtt og 0.3% for meieriprodukt. .

b) Konstruer ein kontrast for å teste hypotesa at den forventa prisauken for kjøtt er den same som den gjennomsnittlege forventa prisauken for dei 4 andre kategoriane. Utfør testen. Kva blir konklusjonen med eit signifikansnivå valgt til 0.05?

Alle tal i variansanalysetabellen utrekna fra dei observerte dataane, kunne vore konstruert med informasjon om verdiane for $\bar{y}_{.j}, j = 1, 2, \ldots, 5$ og $\sum_{j=1}^{5} \sum_{i=1}^{10} y_{ij}^2$.

c) Vis at SSTOT kan utreknast dersom denne informasjonen er gitt. Kva blir talet for $\sum_{j=1}^{5} \sum_{i=1}^{10} y_{ij}^2$ for desse dataane?