Viðskipta- og Hagfræðideild
Háskóli Íslands

Tölfræði II, fyrirlestur 4
Helgi Tómasson

Nokkur atriði úr kafla 6.

- Regla að kunna:

  $$f(y, x|\boldsymbol{\phi}) = f(y|x, \boldsymbol{\phi})f(x|\boldsymbol{\phi})$$

  Þessi regla nýtist oft þegar að þarf að skrifa niður þétti fall fyrir háðar hendingar.

- Það nægir að taka lauslega tengslahugtökin á bls 271-272.
- Lesið vel um fylgnistuðulinn (correlation).

Kafli 7.

- Þessi kafli gengur út á að tjá sig um Y gefið $X = x$.
- Grundvallaratriði eru fyrstu tvö mómentin:

  $$E(Y|X = x) \quad \text{og} \quad V(Y|X = x)$$

- Margvíð normaldreifing grundvallaratriði hér (sjá bls. 342). Hin dæmin í kaflanum eru sýnidæmi til að gefa dæmi um eitthvað annað. Ég tel að þau dæmi séu lítið notuð í praxís. Áhugasamir ættu að kynna sér „weak exogeneity” hugtakið í kafla 7.4.

Stochastic processes, k. 8

- Mikil fræðigrein, hér aðeins tæpt á grunnatriðum.
- Greinin fjallar um hendingar sem eru háðar í tíma.
- Greinin lýsir tilraun þar sem útkoman er ferill:

$$
\begin{aligned}
&X(s,t) \quad s \in S, s \text{ er útkoma úr tilraun,}\\
&S = \text{mengi mögulegra útkoma, sumir skrifa}\\
&X(\omega,t) \quad \omega \in \Omega = \text{ mengi mögulegra útkoma}\\
&\omega \text{ einstök útkoma}
\end{aligned}
$$

- $X(.,.)$ er fall sem úthlutar parinu útkomu og tímapunkti gildi. Fyrir gefna útkomu s (eða ω) má teikna feril.
- Mikið af gögnum í hagfræði og fjármálum kemur úr kerfi þar sem mælingar eru háðar í tíma. Tilgangurinn með að kynna þessa líkindafræði hér er náttúrulega að skapa meðvitund sem nýtist við greiningu á gögnum.
- Takið eftir flokkunarkerfinu í mynd 8.4. Flokkunin byggir m.a. á a) hvernig tíminn er mældur og b) hvort X er samfelld eða sundurslitið.
- Grundvallaratriði er að X gildin á einum tímapunkti geta verið háð X gildum á öðrum tímapunktum (dependency hugtak í bók). Annað grundvallaratriði er að dreifing á einum tímapunkti getur verið öðruvísi en dreifing á öðrum tímapunkti (heterogeneity).

- Nokkur grundvallar „dependency hugtök" (sjá bls. 420 í bók):

  Óháð á öllum tímapunktum

  Markov tengsl = nýliðin fortíð skiptir máli, t.d.

  $f(X(t)|X(t-1), X(t-2), \ldots) = f(X(t)|X(t-1))$

  non-correlation $Corr(X(t), X(k)) = 0$ ef $t \neq k$

  Martingale difference sequence $E(X(t)|\text{fortíð}) = 0$

- Það má lesa lauslega um *mixing conditions* , nema *ergodicity*. Þetta er líkindafræði hugtak sem við ekki notum. Það fjallar um hvernig tengsl skuli fjara út með vaxandi tíma (afar ógætilega orðað).

- Ferli $X(t)$ er *ergodic* ef X gildi sem er langt frá hvert öðru í tíma eru „næstum" óháð. Þetta hugtak er nauðsynlegt ef gagnasöfnun á tímaröðum (tekið fyrir síðar) á að hafa merkingu.

- Nokkur atriði um „heterogeneity/homogeneity" (dreifingar) hugtök.
- Einsdreifðar, $f(X(t) =$ ekki fall af t.
- Sístæðni (stationarity).

  $f(x(t_1), \ldots, x(t_n)) = f(x(t_1 + h), \ldots, x(t_n + h))$
  dreifing óháð hliðrun í tima$= =$ *strict stationarity*
  $E(X(t)) = \mu$ ekki fall af tíma og
  $\gamma(k) = E(X(t) - \mu)(X(t - k) - \mu)$ ekki fall af tíma, bara k
  $=$ *weak stationarity* $=$ *second order stationarity*

  $\gamma(k)$ er kallað „auto-covariance" fall og $\gamma(k)/\gamma(0)$ „auto-correlation" (sjálffylgni) fall.
- *Exchangeability*=háðar en röð skiptir ekki máli.

Nokkur dæmi um ferli

- IID normal hendingar=random úrtak
- *White-noise*:

  $$E(X(t)) = 0, \quad V(X(t)) = \sigma^2 \quad E(X(t)X(s)) = 0 \text{ ef } t \neq s$$

  normal white-noise ef $X(t)$ er normaldreifð.
  Normal white-noise= random úrtak úr normaldrefingu.
- *Random-walk*: Ef $X(t), t = 1, 2, 3 \ldots$ er IID þá er:

  $$Y(t) = \sum_{j=1}^{t} X(t)$$

  random-walk.
- *Martingale*: $E(Y(t)|Y(t-1)) = Y(t-1)$ og $E(|Y(t)|) < \infty$

Mikilvæg ferli sem byggja á normaldreifingu

+

- AR(1) ferli:

$$\varepsilon_t \text{ er normal white-noise } V(\varepsilon_t) = \sigma^2$$
$$X_t = \alpha X_{t-1} + \varepsilon_t$$

- Ef krafist er að X_t sé stationary, þá fæst:

$$E(X_t) = \alpha E(X_{t-1}) \text{ þá fæst } E(X_t) = 0$$
$$V(X_t) = \alpha^2 V(X_t) \text{ þá fæst } V(X_{t-1}) = \frac{\sigma^2}{1-\alpha^2}$$

  því er nauðsynlegt að $|\alpha| < 1$.

- Áhugavert tilfelli af AR(1): $\alpha = 1$. Þá er ferlið ekki sístætt.

$$E(X_t|X_0) = X_0 \quad V(X_t) = t\sigma^2$$

  Þetta er normal random-walk.

- Hvað skeður ef random-walk er mældur þéttar og þéttar?
- Hermum random-walk, fyrir $t = 1, \ldots, T$ gefum okkur $X_0 = 0$ og teiknum ferilinn X á móti t/T. Endurtökum fyrir stærra og stærra T. Tíminn er mældur á „næstum" samfelldum kvarða á bilinu (0,1).
- Það sem út kemur líkist *Brown*-hreyfingu, $B(t)$. Einnig oft kallað Wiener-ferli, $W(t)$.

Eiginleikar Brown-hreyfingar

- Samfelldur ferill kemur út. X er samfelldur, t er mælt á samfelldum kvarða.
- Ferlið er minnislaust. Þ.e. hreyfing á einu tímabili hefur ekki áhrif á hreyfing á öðru sundurlægu tímabili.

- Þ.e.:

$$\begin{aligned}
&E(X(t)) = X(0) \quad X(0) \text{ gefið byrjunargildi}\\
&V(X(t)) = \sigma^2 t\\
&X(t_4) - X(t_3) \text{ óháð } X(t_2) - X(t_1) \text{ ef } t_1 < t_2 < t_3 < t_4\\
&X(t) \sim N(X(0), \sigma^2 t)
\end{aligned}$$

- $B(t)$ er kennt við Brown 19-aldar, líffræðing og $W(t)$ við Wiener sem sannað tilvist ferlisins á afar erfiðan hátt . Tilvist hægt að sanna miklu auðveldar með Kolmogorov extension theorem.

- Brown, Einstein og Bachelier unnu með hugtakið.

- Ýmis ferli er leidd út frá Brown-hreyfingu, Ornstein-Uhlenbeck, Brownian-bridge, geometric-Brownian-motion(GBM). Þeir sem fara í fjármál eiga eftir að nota GBM(Black-Scholes), CIR (Cox-Ingersoll-Ross), CEV(Constant-elasticity-volatility).

Poisson ferli

- Poisson ferli er talningarferli. Það lýsir fjölda slysa eða viðlíka atburða sem falli af tíma.
- Poisson ferli er minnislaust (eins og Brown-hreyfingin), þ.e. að fjöldi slysa á einu tímabili hefur ekki áhrif á önnur.
- Biðtími eftir næsta slysi er exponential dreifður. Skoðið mynd 8.8 í bók.
- Væntanlegt gildi og varíans fyrir Poisson ferli er:

  $$E(N(t)) = \lambda t \quad V(N(t)) = \lambda t$$

  (berið saman við Brown-hreyfingu).

AR, MA og ARMA ferli

- Normal MA ferli:

  $$\begin{aligned} &\varepsilon_t \quad \text{er normal white noise} \\ &X_t = \varepsilon_t + a_1\varepsilon_{t-1} + \cdots a_m\varepsilon_{t-m} \end{aligned}$$

  er kallaður MA(m).
- AR(1) má útvíkka í AR(p):

  $$\begin{aligned} &\varepsilon_t \quad \text{er normal white noise} \\ &X_t = \alpha_1 X_{t-1} + \cdots + \alpha_p X_{t-p} + \varepsilon_t \end{aligned}$$

- Það er hugsanlegt að blanda saman AR og MA liðum og fá ARMA(p,m):

  $$\begin{aligned} &\varepsilon_t \quad \text{er normal white noise} \\ &X_t = \alpha_1 X_{t-1} + \cdots + \alpha_p X_{t-p} + \varepsilon_t + a_1\varepsilon_{t-1} + \cdots a_m\varepsilon_{t-m} \end{aligned}$$

  ARMA líkön eru mikið notuð í hagnýtri tímaraðagreiningu (t.d. Box-Jenkins aðferðir).

Samleitnihugtök, kafli 9. Aðallega 9.6 og 9.9.

- Erfiður kafli með nýjum hugtökum.
- Úr tölfræði I (Newbold) þekkið þið nálgun á binomial dreifingu með poisson dreifingu og nálgun á úrtaksdreifingum með normaldreifingu. Til að geta gert þetta þarf að nota tiltekin samleitnihugtök (convergence mode)
- Lögmál stórra talna, weak og strong, hvað þýðir þetta?
- Byrja á kafla 9.9, modes of convergence. Hvað þýðir $X_n \underset{n\to\infty}{\longrightarrow}$?
- Skilgreining 1: *Pointwise convergence*

  $$X_n(s) \underset{n\to\infty}{\longrightarrow} X(s) \quad \text{fyrir öll } s.$$

  Auðskilið hugtak en of strangt til að vera praktískt.
- Skilgreining 2: *Almost sure convergence*:

  $$X_n \xrightarrow[n\to\infty]{a.s.} X \Leftrightarrow \lim_{n\to\infty} P(\bigcup_{k>n} |X_k(s) - X(s)| > \varepsilon) = 0$$
- Skilgreining 3: *Convergence in probability*:

  $$X_n \xrightarrow[n\to\infty]{p} X \Leftrightarrow \lim_{n\to\infty} P(|X_n(s) - X(s)| > \varepsilon) = 0$$

  einnig skrifað

  $plim(X_n) = X$
- Skilgreining 4: *Convergence in distribution* :

  $$X_n \xrightarrow[n\to\infty]{d} X \Leftrightarrow \lim_{n\to\infty} F_{X_n}(x) = F_X(x)$$

  í öllum punktum þar sem F_X er samfellt
- Skilgreining 5: *Mean-square convergence*:

  $$X_n \xrightarrow[n\to\infty]{m.s.} \Leftrightarrow X \lim_{n\to\infty} E(|X_n - X|^2) = 0$$
- Skoðið tengsl samleitnihugtakanna í lemmu 5 á bls. 509.
- Samleitni fyrir runur af hendingum er í eðli sínu samleitni fyrir runur af föllum.

Central limit theorem

- Ef gefið er random úrtak $X_1, \ldots, X_n$, þ.a. $E(X_i) = \mu$ og $V(X_i) = \sigma^2$. Ath. annað er ekki gefið. Þá gildir að:

$$
\begin{aligned}
&S_n = X_1 + \cdots + X_n, \\
&S_n \xrightarrow[n\to\infty]{d} N(n\mu, n\sigma^2) \quad \text{þ.e.} \\
&\frac{S_n - n\mu}{\sigma\sqrt{n}} \quad \text{er nokkurn veginn } N(0,1)
\end{aligned}
$$

- Þetta má sanna með því að nota Taylor útvíkkun á characteristicfalli fyrir S_n. Gangurinn er nokkurn veginn:

  $$\phi_{S_n}(t) = \phi_{X_1+\cdots+X_n}(t) = E(exp(it(X_1 + \cdots + X_n)) =$$
  $$E(exp(itX_1))\dots E(exp(itX_n)) = \phi_{X_1}(t)\cdots\phi_{X_n}(t) = [\phi_X(t)]^n$$

  nú er $[\phi_X(t)]^n$ Taylor útvíkkað í 0

  Taylor útvíkkun á falli f í punkti a er

  $$f(a)(x-0)^0 + f'(a)(x-a) + \frac{1}{2}f''(a)(x-a)^2 + \cdots$$

  það er jafngott og e.t.v. einfaldara að Taylor útvíkka $log(\phi_{S_n}(t))$

  $$\log(\phi_{S_n}(0)) + \frac{\phi'_{S_n}(0)t}{\phi_{S_n}(0)} + \left(\frac{\phi''_{S_n}(0)}{2\phi_{S_n}(0)} - \frac{\phi'_{S_n}(0)^2}{2\phi_{S_n}(0)^2}\right)t^2 + \cdots$$

  Nú þarf að nota að:

  $$\phi_{S_n}(0) = 1 \quad \phi'_{S_n}(0) = in\mu \quad \phi''_{S_n}(0) = -n\sigma^2 - n^2\mu^2$$

  þ.e.(ath. $i^2 = -1$.

  $$log(\phi_{S_n}(t) = in\mu t - n\sigma^2 t/2 + \text{ rest-liður}$$

  Characterisic-fall fyrir normaldreifingu er:

  $$exp(i\mu^* t - \sigma^* t^2/2)$$

  Þar sem:

  $$\phi_{S_n}(t) \simeq exp(in\mu t - n\sigma^2 t/2)$$

  er drefing S_n nokkurn veginn normal, skv. theorem bls. 113.

Hvað er framundan

- Höfum tekið grófa yfirferð í líkindafræði
- Næst er hagnýting líkindafræðinnar í tölfræði=ályktunarfræði.
- Þurfum að gefa líkum einhverja merkingu. Hér er áherslan á að túlka líkur sem mælikvarði á tíðni. Önnur nálgun er að túlka líkur sem mælikvarða á vissu (Baeysísk túlkun).
- Kafli 10 er eins konar brú frá líkindafræði yfir í ályktunarfræði. Berið saman við Newbold.
- Kaflar 11-15 í Spanos fjalla um uppbyggingu ályktunarfræði