<table>
<tr><td>Háskóli Íslands</td><td>09.12.21
Grundvöllur tölfræðinnar</td><td>Raunvísindadeild</td></tr>
<tr><td>Fimmtudagur</td><td>3. desember 2009</td><td>kl 09:00-12:00</td></tr>
</table>

Leyfileg hjálpargögn: Skriffæri og reiknivél. Jöfnublöð eru aftast í prófinu.		Dæmin vega jafnt.

Munið að eyða ekki miklum tíma í að reyna að sanna t.d. fullkomnleika. Betra er að byrja á að fullyrða slíkt og sanna síðar ef tími gefst til.

1: Látið $X_1, \ldots, X_n \sim n(\xi, \sigma^2)$ og $Y_1, \ldots, Y_m \sim n(\eta, \tau^2)$ vera óháð úrtök úr tveimur normaldreifingum, með nánari forsendum eins og lýst er í hverjum lið.

(a) Finnið ósmækkanlega nægjanlega hendingu ef gefið er að $\sigma^2 = \tau^2$.

(b) Finnið sennileikametlana fyrir ξ, η og σ^2ef gefið er að $\sigma^2 = \tau^2$.

(c) Finnið sennileikaprófið (LRT) fyrir $H_0 : \sigma^2 = \tau^2$ gegn $H_1 : \sigma^2 \neq \tau^2$.

(d) Lítið nánar á fyrra hendingasafnið, $X_1, \ldots, X_n \sim n(\xi, \sigma^2)$ (óháðar). Sýnið að

$$\bar{X}^2 - \frac{S^2}{n(n-1)}$$

er bestur óbjagaðra metla (UMVUE) fyrir ξ^2.

[Leiðbeiningar: Munið að einfalda má jöfnurnar með $a_1, \ldots, a_n$ og b eru tölur þá gildir $\sum (a_i - b)^2 = \sum (a_i - \bar{a})^2 + n(\bar{a} - b)^2$]

Lítum í dæmum 2-3 á óháðar, einsdreifðar hendingar $X_1, \ldots, X_n$, sem lúta geometrískri dreifingu, þ.e. $P[X_i = x] = p(1-p)^{x-1}$, fyrir $x = 1,\ 2,\ \ldots$.

2: (a) Finnið sennileikametilinn, MLE, $\hat{p}$, fyrir p.

(b) Er $\hat{p}$ mótsagnalaus (consistent) fyrir p?

(c) Finnið aðfelludreifingu (asymptotic distribution) fyrir $\hat{p}$.

(d) Er $\hat{p}$ skilvirkur (efficient)?

[Leiðbeiningar: Í (b) og (c) getur verið gott að byrja á að hugsa um aðfelludreifingu $\bar{X}$ og síðan hvernig aðfelludreifing $\hat{p}$ tengist þeirri niðurstöðu.]

3: Geometríska dreifingin, eins og hún er sett fram hér, tilsvarar líkum á því að fá jákvæða niðurstöðu eftir x óháðar tilraunir. Umritið hana þannig að x tákni fjölda tilrauna áður en fyrst kemur jákvæð niðurstaða, svo $P[X_i = x] = p(1-p)^x$, fyrir $x = 0,\ 1,\ 2,\ \ldots$.

(a) Finnið spanfall vægis fyrir þessa breyttu framsetningu og beitið því til að sjá hver líkindadreifing $\sum_i X_i$ verður.

(b) Finnið sterkasta prófið (UMP) fyrir $H_0 : p = p_0$ gegn $H_a : p > p_0$.

4: Lítið á n óháðar hendingar sem lúta jafndreifingu á $[\theta - 1/2, \theta + 1/2]$, þ.e. $X_1, \ldots, X_n \sim U(\theta - 1/2, \theta + 1/2)$.

(a) Finnið ósmækkanlega nægjanlega hendingu og sýnið að hún er ekki fullkomin (complete).

(b) Veljið á rökstuddan hátt “heppilega” vendihendingu (pivotal quantity) og umsnúið henni til að fá $100(1-\alpha)\%$ öryggisbil fyrir θ.

(c) Sýnið, hvernig velja má “löglega” endapunkta og rökstyðjið síðan hvernig má gera bilið sem styst.

[Munið að aðalatriðið er að lýsa aðferðunum.]

Aukadæmi: Í stað hefðbundinnar forsendu um að niðurstaða (X) fjöldatalningar lúti tvíkostadreifingu er stundum bætt við forsendu um að líkurnar p í hverju kasti lúti beta-dreifingu. Finnið massafall, væntigildi og dreifni X ef $X|p \sim b(n,p)$ og $p \sim B(\alpha, \beta)$.

[Leiðbeiningar: Gefið ykkur það sem þarf um betadreifinguna og tvíkostadreifinguna. Athugið að auðveldara getur verið að nota V[X]=V[E[X|p]]+E[V[X|p]] heldur en massafallið þegar kemur að dreifninni (og tilsvarandi fyrir væntigildið).]

Til minnis

$$W = t(\mathbf{X})$$

MSE
$$E_\theta\left[(W-\theta)^2\right] = E_\theta\left[(W - E_\theta[W])^2\right] + (E_\theta[W]-\theta)^2$$

CR1
$$V_\theta[W] \geq \frac{\left(\frac{d}{d\theta}E_\theta[W]\right)^2}{E_\theta\left[\left(\frac{\partial}{\partial\theta} ln f_\theta(\mathbf{X})\right)^2\right]}$$

CR2
$$V_\theta[W] \geq \frac{1}{-nE_\theta\left[\frac{\partial^2}{\partial\theta^2} ln f_\theta(X_1)\right]}$$

LRT
$$\lambda(\mathbf{x}) = \frac{\sup_{\Theta_0} L_{\mathbf{x}}(\theta)}{\sup_{\Theta} L_{\mathbf{x}}(\theta)}$$

NP
$$f_{\theta_1}(\mathbf{x}) > k f_{\theta_0}(\mathbf{x})$$

$$\beta(\theta) = P_\theta[\phi(\mathbf{X}) = 1]$$

$$\hat{\beta} = \frac{\Sigma(x-\bar{x})(y-\bar{y})}{\Sigma(x-\bar{x})^2}$$

$$V\left[\hat{\beta}\right] = \frac{\sigma^2}{\Sigma(x-\bar{x})^2}$$

$$V_\theta\left[h\left(\hat{\theta}\right)\right] \simeq \frac{\left[h'(\hat{\theta})\right]^2}{-\left.\frac{\partial^2}{\partial\theta^2} ln L_{\mathbf{x}}(\theta)\right|_{\theta=\hat{\theta}}}$$

MLR
$$\exists\ t:\mathbb{R}\to\mathbb{R} \text{ þ.a. } g(x) := \frac{f_{\theta'}(x)}{f_\theta(x)} \text{ er vaxandi í } t(x) \text{ ef } \theta < \theta'$$

Δ: $\sqrt{n}\left(X_n - \theta\right) \to n\left(0,\sigma^2\right) \Rightarrow \sqrt{n}\left(g(X_n) - g(\theta)\right) \to n\left(0, \sigma^2\left(g'(\theta)\right)^2\right)$

Table of Common Distributions

Discrete Distributions

Bernoulli(p)

pmf	$P(X=x\|p)=p^x(1-p)^{1-x};\quad x=0,1;\quad 0\le p\le 1$
mean and variance	$\mathrm{E}X=p,\quad \mathrm{Var}\,X=p(1-p)$
mgf	$M_X(t)=(1-p)+pe^t$

Binomial(n,p)

pmf	$P(X=x\|n,p)=\binom{n}{x}p^x(1-p)^{n-x};\quad x=0,1,2,\ldots,n;\quad 0\le p\le 1$
mean and variance	$\mathrm{E}X=np,\quad \mathrm{Var}\,X=np(1-p)$
mgf	$M_X(t)=[pe^t+(1-p)]^n$
notes	Related to Binomial Theorem (Theorem 3.2.2). The *multinomial* distribution (Definition 4.6.2) is a multivariate version of the binomial distribution.

Discrete uniform

pmf	$P(X=x\|N)=\frac{1}{N};\quad x=1,2,\ldots,N;\quad N=1,2,\ldots$
mean and variance	$\mathrm{E}X=\frac{N+1}{2},\quad \mathrm{Var}\,X=\frac{(N+1)(N-1)}{12}$
mgf	$M_X(t)=\frac{1}{N}\sum_{i=1}^{N}e^{it}$

Geometric(p)

pmf	$P(X=x\|p)=p(1-p)^{x-1};\quad x=1,2,\ldots;\quad 0\le p\le 1$
mean and variance	$\mathrm{E}X=\frac{1}{p},\quad \mathrm{Var}\,X=\frac{1-p}{p^2}$

mgf	$M_X(t) = \frac{pe^t}{1-(1-p)e^t}, \quad t < -\log(1-p)$
notes	$Y = X - 1$ is negative binomial$(1, p)$. The distribution is *memoryless*: $P(X > s \mid X > t) = P(X > s - t)$.

Hypergeometric

pmf	$P(X = x \mid N, M, K) = \frac{\binom{M}{x}\binom{N-M}{K-x}}{\binom{N}{K}}; \quad x = 0, 1, 2, \ldots, K;$ $M - (N - K) \leq x \leq M; \quad N, M, K \geq 0$
mean and variance	$\mathrm{E}X = \frac{KM}{N}, \quad \mathrm{Var}\, X = \frac{KM}{N}\frac{(N-M)(N-K)}{N(N-1)}$
notes	If $K \ll M$ and N, the range $x = 0, 1, 2, \ldots, K$ will be appropriate.

Negative binomial(r, p)

pmf	$P(X = x \mid r, p) = \binom{r+x-1}{x} p^r (1-p)^x; \quad x = 0, 1, \ldots; \quad 0 \leq p \leq 1$
mean and variance	$\mathrm{E}X = \frac{r(1-p)}{p}, \quad \mathrm{Var}\, X = \frac{r(1-p)}{p^2}$
mgf	$M_X(t) = \left(\frac{p}{1-(1-p)e^t}\right)^r, \quad t < -\log(1-p)$
notes	An alternate form of the pmf is given by $P(Y = y \mid r, p) = \binom{y-1}{r-1} p^r (1-p)^{y-r}$, $y = r, r+1, \ldots$. The random variable $Y = X + r$. The negative binomial can be derived as a gamma mixture of Poissons. (See Exercise 4.34.)

Poisson(λ)

pmf	$P(X = x \mid \lambda) = \frac{e^{-\lambda}\lambda^x}{x!}; \quad x = 0, 1, \ldots; \quad 0 \leq \lambda < \infty$
mean and variance	$\mathrm{E}X = \lambda, \quad \mathrm{Var}\, X = \lambda$
mgf	$M_X(t) = e^{\lambda(e^t - 1)}$

Continuous Distributions

***Beta*(α, β)**

pdf	$f(x\|\alpha,\beta) = \frac{1}{B(\alpha,\beta)} x^{\alpha-1}(1-x)^{\beta-1}, \quad 0 \le x \le 1, \quad \alpha > 0, \quad \beta > 0$
mean and variance	$\mathrm{E}X = \frac{\alpha}{\alpha+\beta}, \quad \mathrm{Var}\, X = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$
mgf	$M_X(t) = 1 + \sum_{k=1}^{\infty} \left(\prod_{r=0}^{k-1} \frac{\alpha+r}{\alpha+\beta+r} \right) \frac{t^k}{k!}$
notes	The constant in the beta pdf can be defined in terms of gamma functions, $B(\alpha,\beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$. Equation (3.2.18) gives a general expression for the moments.

***Cauchy*(θ, σ)**

pdf	$f(x\|\theta,\sigma) = \frac{1}{\pi\sigma} \frac{1}{1+\left(\frac{x-\theta}{\sigma}\right)^2}, \quad -\infty < x < \infty; \quad -\infty < \theta < \infty, \quad \sigma > 0$
mean and variance	do not exist
mgf	does not exist
notes	Special case of Student's t, when degrees of freedom = 1. Also, if X and Y are independent $\mathrm{n}(0,1)$, X/Y is Cauchy.

***Chi squared*(p)**

pdf	$f(x\|p) = \frac{1}{\Gamma(p/2)2^{p/2}} x^{(p/2)-1} e^{-x/2}; \quad 0 \le x < \infty; \quad p = 1, 2, \ldots$
mean and variance	$\mathrm{E}X = p, \quad \mathrm{Var}\, X = 2p$
mgf	$M_X(t) = \left(\frac{1}{1-2t}\right)^{p/2}, \quad t < \frac{1}{2}$
notes	Special case of the gamma distribution.

***Double exponential*(μ, σ)**

pdf	$f(x\|\mu,\sigma) = \frac{1}{2\sigma} e^{-\|x-\mu\|/\sigma}, \quad -\infty < x < \infty, \quad -\infty < \mu < \infty, \quad \sigma > 0$
mean and variance	$\mathrm{E}X = \mu, \quad \mathrm{Var}\, X = 2\sigma^2$
mgf	$M_X(t) = \frac{e^{\mu t}}{1-(\sigma t)^2}, \quad \|t\| < \frac{1}{\sigma}$
notes	Also known as the *Laplace* distribution.

Exponential(β)

pdf	$f(x\|\beta) = \frac{1}{\beta}e^{-x/\beta}, \quad 0 \le x < \infty, \quad \beta > 0$
mean and variance	$\mathrm{E}X = \beta, \quad \mathrm{Var}\,X = \beta^2$
mgf	$M_X(t) = \frac{1}{1-\beta t}, \quad t < \frac{1}{\beta}$
notes	Special case of the gamma distribution. Has the *memoryless* property. Has many special cases: $Y = X^{1/\gamma}$ is *Weibull*, $Y = \sqrt{2X/\beta}$ is *Rayleigh*, $Y = \alpha - \gamma \log(X/\beta)$ is *Gumbel*.

F

pdf	$f(x\|\nu_1, \nu_2) = \frac{\Gamma\left(\frac{\nu_1+\nu_2}{2}\right)}{\Gamma\left(\frac{\nu_1}{2}\right)\Gamma\left(\frac{\nu_2}{2}\right)} \left(\frac{\nu_1}{\nu_2}\right)^{\nu_1/2} \frac{x^{(\nu_1-2)/2}}{\left(1+\left(\frac{\nu_1}{\nu_2}\right)x\right)^{(\nu_1+\nu_2)/2}};$ $0 \le x < \infty; \quad \nu_1, \nu_2 = 1, \ldots$
mean and variance	$\mathrm{E}X = \frac{\nu_2}{\nu_2-2}, \quad \nu_2 > 2,$ $\mathrm{Var}\,X = 2\left(\frac{\nu_2}{\nu_2-2}\right)^2 \frac{(\nu_1+\nu_2-2)}{\nu_1(\nu_2-4)}, \quad \nu_2 > 4$
moments (mgf does not exist)	$\mathrm{E}X^n = \frac{\Gamma\left(\frac{\nu_1+2n}{2}\right)\Gamma\left(\frac{\nu_2-2n}{2}\right)}{\Gamma\left(\frac{\nu_1}{2}\right)\Gamma\left(\frac{\nu_2}{2}\right)} \left(\frac{\nu_2}{\nu_1}\right)^n, \quad n < \frac{\nu_2}{2}$
notes	Related to chi squared ($F_{\nu_1,\nu_2} = \left(\frac{\chi^2_{\nu_1}}{\nu_1}\right) / \left(\frac{\chi^2_{\nu_2}}{\nu_2}\right)$, where the χ^2s are independent) and t ($F_{1,\nu} = t^2_\nu$).

Gamma(α, β)

pdf	$f(x\|\alpha, \beta) = \frac{1}{\Gamma(\alpha)\beta^\alpha}x^{\alpha-1}e^{-x/\beta}, \quad 0 \le x < \infty, \quad \alpha, \beta > 0$
mean and variance	$\mathrm{E}X = \alpha\beta, \quad \mathrm{Var}\,X = \alpha\beta^2$
mgf	$M_X(t) = \left(\frac{1}{1-\beta t}\right)^\alpha, \quad t < \frac{1}{\beta}$
notes	Some special cases are exponential ($\alpha = 1$) and chi squared ($\alpha = p/2$, $\beta = 2$). If $\alpha = \frac{3}{2}$, $Y = \sqrt{X/\beta}$ is *Maxwell*. $Y = 1/X$ has the *inverted gamma distribution*. Can also be related to the Poisson (Example 3.2.1).

Logistic(μ, β)

pdf	$f(x\|\mu, \beta) = \frac{1}{\beta} \frac{e^{-(x-\mu)/\beta}}{[1+e^{-(x-\mu)/\beta}]^2}, \quad -\infty < x < \infty, \quad -\infty < \mu < \infty, \quad \beta > 0$
mean and variance	$\mathrm{E}X = \mu, \quad \mathrm{Var}\,X = \frac{\pi^2\beta^2}{3}$

mgf	$M_X(t) = e^{\mu t}\Gamma(1-\beta t)\Gamma(1+\beta t), \quad \lvert t\rvert < \frac{1}{\beta}$
notes	The cdf is given by $F(x\lvert\mu,\beta) = \frac{1}{1+e^{-(x-\mu)/\beta}}$.

Lognormal(μ, σ^2)

pdf	$f(x\lvert\mu,\sigma^2) = \frac{1}{\sqrt{2\pi}\sigma}\frac{e^{-(\log x-\mu)^2/(2\sigma^2)}}{x}, \quad 0 \le x < \infty, \quad -\infty < \mu < \infty, \quad \sigma > 0$
mean and variance	$\mathrm{E}X = e^{\mu+(\sigma^2/2)}, \quad \mathrm{Var}\,X = e^{2(\mu+\sigma^2)} - e^{2\mu+\sigma^2}$
moments (mgf does not exist)	$\mathrm{E}X^n = e^{n\mu+n^2\sigma^2/2}$
notes	Example 2.3.5 gives another distribution with the same moments.

Normal(μ, σ^2)

pdf	$f(x\lvert\mu,\sigma^2) = \frac{1}{\sqrt{2\pi}\sigma}e^{-(x-\mu)^2/(2\sigma^2)}, \quad -\infty < x < \infty, \quad -\infty < \mu < \infty, \quad \sigma > 0$
mean and variance	$\mathrm{E}X = \mu, \quad \mathrm{Var}\,X = \sigma^2$
mgf	$M_X(t) = e^{\mu t+\sigma^2t^2/2}$
notes	Sometimes called the *Gaussian* distribution.

Pareto(α, β)

pdf	$f(x\lvert\alpha,\beta) = \frac{\beta\alpha^\beta}{x^{\beta+1}}, \quad a < x < \infty, \quad \alpha > 0, \quad \beta > 0$
mean and variance	$\mathrm{E}X = \frac{\beta\alpha}{\beta-1}, \quad \beta > 1, \quad \mathrm{Var}\,X = \frac{\beta\alpha^2}{(\beta-1)^2(\beta-2)}, \quad \beta > 2$
mgf	does not exist

---

t

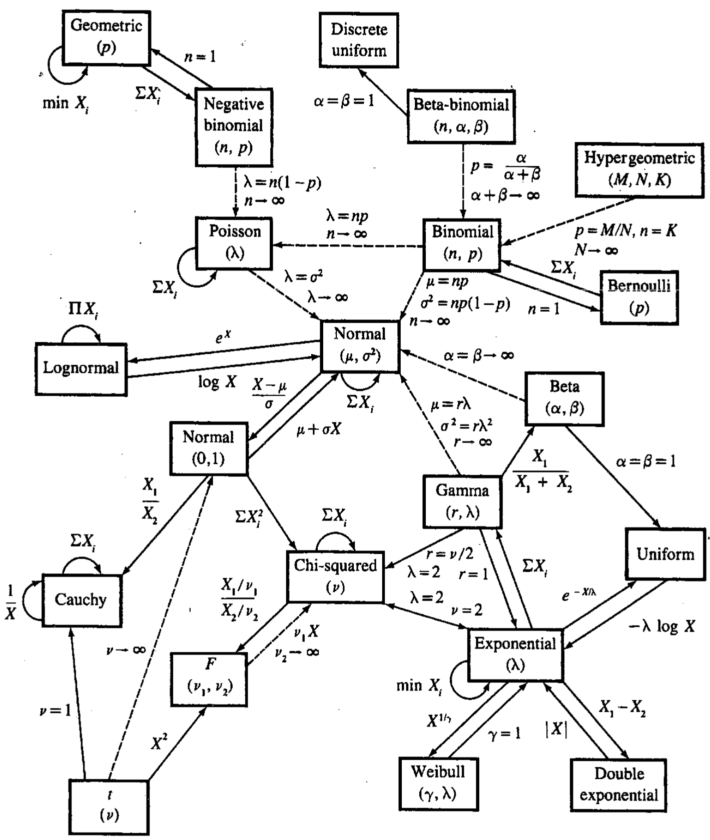
pdf	$f(x\lvert\nu) = \frac{\Gamma(\frac{\nu+1}{2})}{\Gamma(\frac{\nu}{2})}\frac{1}{\sqrt{\nu\pi}}\frac{1}{\left(1+\left(\frac{x^2}{\nu}\right)\right)^{(\nu+1)/2}}, \quad -\infty < x < \infty, \quad \nu = 1, \ldots$
mean and variance	$\mathrm{E}X = 0, \quad \nu > 1, \quad \mathrm{Var}\,X = \frac{\nu}{\nu-2}, \quad \nu > 2$
moments (mgf does not exist)	$\mathrm{E}X^n = \frac{\Gamma(\frac{n+1}{2})\Gamma(\frac{\nu-n}{2})}{\sqrt{\pi}\Gamma(\frac{\nu}{2})}\nu^{n/2}$ if $n < \nu$ and even, $\mathrm{E}X^n = 0$ if $n < \nu$ and odd.
notes	Related to F ($F_{1,\nu} = t_\nu^2$).

Uniform(a, b)

pdf	$f(x\|a,b) = \frac{1}{b-a}, \quad a \le x \le b$
mean and variance	$\mathrm{E}X = \frac{b+a}{2}, \quad \mathrm{Var}\, X = \frac{(b-a)^2}{12}$
mgf	$M_X(t) = \frac{e^{bt}-e^{at}}{(b-a)t}$
notes	If $a = 0$ and $b = 1$, this is a special case of the beta ($\alpha = \beta = 1$).

Weibull(γ, β)

pdf	$f(x\|\gamma,\beta) = \frac{\gamma}{\beta} x^{\gamma-1} e^{-x^\gamma/\beta}, \quad 0 \le x < \infty, \quad \gamma > 0, \quad \beta > 0$
mean and variance	$\mathrm{E}X = \beta^{1/\gamma}\Gamma\left(1+\frac{1}{\gamma}\right), \quad \mathrm{Var}\, X = \beta^{2/\gamma}\left[\Gamma\left(1+\frac{2}{\gamma}\right) - \Gamma^2\left(1+\frac{1}{\gamma}\right)\right]$
moments	$\mathrm{E}X^n = \beta^{n/\gamma}\Gamma\left(1+\frac{n}{\gamma}\right)$
notes	The mgf exists only for $\gamma \ge 1$. Its form is not very useful. A special case is exponential ($\gamma = 1$).



Relationships among common distributions. Solid lines represent transformations and special cases, dashed lines represent limits. Adapted from Leemis (1986).