

Sannolikhet och statistik A1

Johan Jonasson

April 2019

Transformerering av sv. Uppgift: Du har en sv X med förd F_1 . Finn en fkn g så att $g(X)$ har förd F_2 .

Proposition: Om F_1 och F_2 är inverterbara och $X \sim F_1$ gäller att $Y = F_2^{-1}(F_1(X)) \sim F_2$.

Bevis:

$$\mathbb{P}(Y \leq y) = \mathbb{P}(F_1(X) \leq F_2(y)) = \mathbb{P}(X \leq F_1^{-1}(F_2(y))) = F_2(y).$$

Specialfall: Om $U \sim \text{likf}(0, 1)$ gäller att $F^{-1}(U) \sim F$. □

Exempel: Använd en $\text{likf}(0, 1)$ -fördelad sv U till att simulera en $\text{exp}(\lambda)$ -fördelning. Det gäller att $F(x) = 1 - e^{-\lambda x}$. Alltså

$$F^{-1}(u) = -\frac{\ln(1 - u)}{\lambda}.$$

Alltså är $-\ln(1 - U)/\lambda \sim \text{exp}(\lambda)$. □

Minns att $X_1, \dots, X_n$ oberoende och $\exp(\lambda)$ medför $\sum_1^n X_k \sim \Gamma(n, \lambda)$. Vi hade

$$F(x) = 1 - e^{-\lambda x} \sum_{k=0}^{n-1} \frac{(\lambda x)^k}{k!}.$$

Derivering ger

$$f(x) = \frac{1}{(n-1)!} \lambda^n x^{n-1} e^{-\lambda x}.$$

Mer allmänt säger man för parametrar $\alpha, \lambda > 0$ att $X \sim \Gamma(\alpha, \lambda)$ om

$$f(x) = \frac{1}{C} \lambda^\alpha x^{\alpha-1} e^{-\lambda x}.$$

χ^2 -fördelning. Om $Z_1, Z_2, \dots, Z_n$ är oberoende och $N(0, 1)$ -fördelade säger man att

$$X = \sum_{k=1}^n Z_k^2$$

är χ^2 -fördelad med n frihetsgrader. Vi får

$$\mathbb{E}[X] = n, \text{Var}(X) = 2n.$$

Intressant faktum:

$$\begin{aligned}
\mathbb{P}(Z_1^2 + Z_2^2 \leq x) &= \mathbb{P}((Z_1, Z_2) \in B(0, \sqrt{x})) \\
&= \frac{1}{2\pi} \iint_{B(0, \sqrt{x})} e^{-\frac{1}{2}(t_1^2 + t_2^2)} dt_1 dt_2 \\
&= \frac{1}{2\pi} \int_0^{2\pi} \int_0^{\sqrt{x}} r e^{-\frac{1}{2}r^2} dr d\theta \\
&= [-e^{-\frac{1}{2}r^2}]_0^{\sqrt{x}} = 1 - e^{-\frac{1}{2}x}.
\end{aligned}$$

Alltså: $Z_1^2 + Z_2^2 \sim \exp(1/2)$, dvs $\chi_2^2 = \exp(1/2)$. Det följer att att om n är jämnt gäller

$$\sum_{k=1}^n Z_k^2 \sim \Gamma(n/2, 1/2).$$

Detta gäller även då n är udda, dvs $\chi_n^2 = \Gamma(n/2, 1/2)$.

Låt $\mathbf{U} = (U_1, \dots, U_n)$ vara oberoende och $N(0, 1)$. Låt M vara en $n \times n$ ortogonalmatrix. Låt $\mathbf{X} = M\mathbf{U}$. Då är $X_1, \dots, X_n$ oberoende och $N(0, 1)$. (Visas inte.) Låt

$$M = [\mathbf{c}_1 \ \mathbf{c}_2 \ \dots \ \mathbf{c}_n]$$

där $\mathbf{c}_1 = [1/\sqrt{n} \ \dots \ 1/\sqrt{n}]^T = (1/\sqrt{n})\mathbf{1}$. Då är

$$\mathbf{X} = \sum_{k=1}^n U_k \mathbf{c}_k.$$

Vi får

$$\bar{X} = \frac{1}{n} \mathbf{1}^T \mathbf{X} = \frac{1}{\sqrt{n}} U_1$$

eftersom $\mathbf{c}_2, \dots, \mathbf{c}_n$ är ortogonala mot $\mathbf{1}$.

Alltså

$$\begin{aligned}
 \sum_{k=1}^n (X_k - \bar{X})^2 &= (\mathbf{X} - \bar{X}\mathbf{1})^T (\mathbf{X} - \bar{X}\mathbf{1}) \\
 &= \left(\mathbf{X} - \frac{1}{\sqrt{n}} U_1 \mathbf{1} \right)^T \left(\mathbf{X} - \frac{1}{\sqrt{n}} U_1 \mathbf{1} \right) \\
 &= \left(\sum_{k=2}^n U_k \mathbf{c}_k \right)^T \left(\sum_{k=2}^n U_k \mathbf{c}_k \right) \\
 &= \sum_{k=2}^n U_k^2.
 \end{aligned}$$

Det följer att $\sum_{k=1}^n (X_k - \bar{X})^2 \sim \chi_{n-1}^2$ och är oberoende av $\bar{X}$, dvs $(n-1)s^2 \sim \chi_{n-1}^2$ och oberoende av $\bar{X}$. Med X_k :na generellt $N(\mu, \sigma^2)$, standardisera och få

$$\frac{(n-1)s^2}{\sigma^2} \sim \chi_{n-1}^2.$$

Konfidensintervall för σ^2 . Låt $X_1, \dots, X_n$ vara sp på $N(\mu, \sigma^2)$. Det gäller att

$$F_{\chi_{n-1}^2}^{-1}(\alpha/2) \leq \frac{(n-1)s^2}{\sigma^2} \leq F_{\chi_{n-1}^2}^{-1}(1 - \alpha/2)$$

med sannol $1 - \alpha$. Vi får

$$\frac{(n-1)s^2}{F_{\chi_{n-1}^2}^{-1}(1 - \alpha/2)} \leq \sigma^2 \leq \frac{(n-1)s^2}{F_{\chi_{n-1}^2}^{-1}(\alpha/2)}$$

med konfidensgrad $1 - \alpha$. Ett uppåt begränsat konfintervall ges av

$$\sigma^2 \leq \frac{(n-1)s^2}{F_{\chi_{n-1}^2}^{-1}(\alpha)}.$$

Felintensitet. Låt T vara en positiv sv. Se som en livstid. Skriv $G(t) = \mathbb{P}(T > t)$. Givet överlevnad till tid t , vad är risken för haveri alldeles strax?

$$\mathbb{P}(T \in (t, t + \Delta t) | T > t) = \frac{\mathbb{P}(T \in (t, t + \Delta t))}{\mathbb{P}(T > t)} \approx \frac{f(t)\Delta t}{G(t)}.$$

Definition: Felintensiteten för T ges av

$$r(t) = \frac{f(t)}{G(t)}.$$

Ofta lättare att specificera och/eller lättare att förstå än f eller G .
Obs att

$$r(t) = -\frac{d}{dt} \ln G(t).$$

Alltså

$$\ln G(t) = - \int_0^t r(s) ds$$

dvs

$$G(t) = e^{-\int_0^t r(s) ds}.$$

Exempel: $T \sim \exp(\lambda) \Leftrightarrow r(t) = \lambda$.

Exempel: $T \sim \text{likf}(0, 1) \Rightarrow$

$$r(t) = \frac{f(t)}{G(t)} = \frac{1}{1-t}, \quad t \in (0, 1).$$

Kraftigt åldrande.

Exempel: Låt

$$G(t) = \frac{1}{(1+t)^2}, \quad t \geq 0.$$

En s.k. tungsvansad fördelning.

$$r(t) = \frac{f(t)}{G(t)} = \frac{2/(1+t)^3}{1/(1+t)^2} = \frac{2}{1+t}.$$

Föryngring.

Exempel: Låt $r(t) = t$, $t \geq 0$. Då blir

$$G(t) = e^{-\int_0^t s \, ds} = e^{-\frac{1}{2}t^2}.$$

(Ganska likt normalförd betingad på att vara större än 0.)