Tillförlitlighetsteori

Gunnar Englund
Matematisk statistik
KTH

Vt 2005

1 Beräkning av moment

Sats: Om $T \geq 0$ har fördelningsfunktionen $F(t),\ t \geq 0$ så gäller om $E(T^r) < \infty$ att

$$E(T^r) = r\int_0^\infty t^{r-1}(1 - F(t))dt$$

$\square$

Bevis: T har tätheten $f(t) = F'(t)$ och

$$E(T^r) = \int_0^\infty t^r f(t) = \text{(partiell integration)} =$$

$$\Big|_0^\infty t^r(F(t) - 1) - \int_0^\infty rt^{r-1}(F(t) - 1)dt.$$

Men eftersom $E(T^r) < \infty$ gäller att

$$\int_A^\infty t^r f(t)dt \to 0 \text{ då } A \to \infty$$

och vi får

$$\int_A^\infty t^r f(t)dt \geq A^r \int_A^\infty f(t)dt = A^r P(T > A).$$

Alltså gäller att $A^r(1 - F(A)) \to 0$ då $A \to \infty$, dvs den utintegrerade biten är 0. $\square$

Bra formler är alltså

$$E(T) = \int_0^\infty (1 - F(t))dt \text{ och } E(T^2) = 2\int_0^\infty t(1 - F(t))dt.$$

Vi betecknar $1 - F(t) = R(t) = P(T > t)$ och $R(t)$ kallas överlevnadsfunktion.

2 Felintensitet

Vi har

$$P(x < T \leq x + h | T > x) = \frac{P(x < T \leq x + h)}{P(T > x)} = \frac{F(x+h) - F(x)}{R(x)} \approx \frac{f(x)h}{R(x)}.$$

Vi definierar därför felintensiteten som denna proportionalitetskonstant.

$$\lambda(x) = \frac{f(x)}{R(x)},$$

och tolkningen är att $P(x < T \leq x + h | T > x) \approx \lambda(x)h$ dvs $\lambda(x)$ är sannolikheten/tidsenhet för fel vid tid x då komponenten klarat sig fram till tiden x.
Eftersom

$$\lambda(x) = \frac{f(x)}{R(x)} = -\frac{d}{dx}\ln(R(x)),$$

så ser vi att

$$R(x) = \exp(-\int_0^x \lambda(u)du) \text{ och } -\ln(R(x)) = \int_0^t \lambda(u)du.$$

För $Exp(\lambda)$-fördelningen har vi tätheten $f(x) = \lambda e^{-\lambda x}$ och överlevnadsfunktionen $R(x) = e^{-\lambda x}$ som ger $\lambda(x) = \lambda$, dvs att felintensiteten är konstant.

Om livslängden är Weibull-fördelad, dvs har överlevnadsfunktionen $R(x) = \exp(-(\lambda x)^c)$ erhålls tätheten $f(x) = c\lambda^c x^{c-1}\exp(-(\lambda x)^c)$ och man får felintensiteten $\lambda(x) = c\lambda^c x^{c-1}$. Denna avtar i x om $0 < c < 1$ och växer i x om $c > 1$. För $c = 1$ är den konstant och detta svarar precis mot $Exp(\lambda)$-fördelningen.

3 Växande och avtagande felintensitet – IFR och DFR

Vi säger att fördelningen har växande felintensitet eller, synonymt, är IFR (Increasing Failure Rate) om felintensiteten $\lambda(x)$ växer i x. På samma sätt definieras DFR (Decreasing Failure Rate) som att $\lambda(x)$ avtar i x.

En lite mer generell definition är att fördelningen är IFR om $-\ln(R(t))$ är konvex i t. Konvexa funktioner ligger under sina kordor, dvs ligger under räta linjer mellan punkter på kurvan.

Tyvärr har inte säkert ett system uppbyggt av oberoende komponenter som var för sig är IFR själv har en växande felintensitet. Vi definierar därför klassen $IFRA$ (A för Average) som att

$$\frac{1}{t}\int_0^t \lambda(u)du \text{ växer i } t.$$

Vi kommer senare att se att system uppbyggda av oberoende $IFRA$-komponenter är självt $IFRA$.

Sats: $IFR \Rightarrow IFRA$

Bevis 1: Antag att $t_1 \leq t_2$. Att fördelningen är IFR betyder alltså att $\lambda(t_1) \leq \lambda(t_2)$. Vi vill visa att

$$\frac{1}{t_1}\int_0^{t_1} \lambda(u)du \leq \frac{1}{t_2}\int_0^{t_2} \lambda(u)du.$$

Vi får

$$\frac{1}{t_2}\int_0^{t_2} \lambda(u)du = (\text{variabelbyte } v = \frac{t_1}{t_2}u) = \frac{1}{t_1}\int_0^{t_1} \lambda(\frac{t_2}{t_1}v)dv \geq$$

$$\geq (\text{ty } t_2/t_1 \geq 1 \text{ och } \lambda(u) \text{ växande}) \geq \frac{1}{t_1}\int_0^{t_1} \lambda(v)dv$$

vilket var vad vi ville visa. □

Bevis 2: Vi utgår från att $-\ln(R(t))$ är konvex. Den börjar i 0 ty $R(0) = 1$ och alltså gäller att (rita figur)

$$-\ln(R(t_1)) \leq \frac{t_1}{t_2}(-\ln(R(t_2)))$$

vilket lätt ger att

$$\frac{1}{t_1}\int_0^{t_1} \lambda(u)du \leq \frac{1}{t_2}\int_0^{t_2} \lambda(u)du$$

eftersom $\ln(R(t)) = -\int_0^t \lambda(u)du$. □

4 Minneslöshet hos exponentialfördelningens

För Exponentialfördelningen $Exp(\lambda)$ gäller att den har felintensiteten $\lambda(t) = \lambda$, dvs den beror ej av t. Exponentialfördelningen är den enda fördelningen som uppfyller detta. Man inser det eftersom fördelningen är bestämd av felintensiteten.

Vi har en intressant minneslöshetsegenskap hos expontenialfördelningen:

$$P(T > x + y | T > x) = \frac{P(T > x + y; T > x)}{P(T > x)} = \frac{P(T > x + y)}{P(T > x)} =$$

$$\frac{e^{-\lambda(x+y)}}{e^{-\lambda x}} = e^{-\lambda y} = P(T > y).$$

Detta kan tolkas som att en komponent med exponentialfördelad livslängd ej åldras – har den överlevt till tiden x är sannolikheten att den klarar ytterligare y tidsenheter samma som sannolikheten att klara y tidsenheter för en ny komponent. Detta hänger naturligtvis ihop med att felintensiteten är konstant. Det är alltså meningslöst att byta en komponent med exponentialfördelad livslängd eftersom den "gamla" komponenten är lika bra som den "nya".

Ett alternativt sätt att uttrycka denna likhet är att utnyttja definitionen av betingad sannolikhet och skriva den $P(T > x + y) = P(T > x)P(T > y)$.

5 Klasserna NBU och NWU

Vi definierar en ny klass av fördelningar NBU (New Better than Used) för fördelningar där man tjänar på att byta ut en använd komponent.

Definition: Fördelningen kallas NBU om $P(T > x+y) \leq P(T > x)P(T > y)$.

Det finns en dual klass NWU (New Worse than Used) om olikheten går åt andra hållet.)

I betingad form skrivs detta $P(T > x + y | T > x) \leq P(T > y)$ dvs med tolkningen att man har större chans att klara ytterligare y tidsenheter om man byter till en ny komponent. Det lönar sig att byta till en ny om man har strikt olikhet!

Om komponentens livslängd är NWU är det dumt att byta eftersom en ny komponent är sämre än den ”gamla”. Det skall naturligtvis inte tolkas som att komponenten egentligen blir bättre med tiden, utan att populationen från början består av dåliga respektive bra enheter. Efter hand som tiden går utan komponentfel blir sannolikheten större för att den komponent vi studerar verkligen är en av de bra. Sätter vi in en ny, finns risken att den är av den dåliga typen.

Sats: $IFRA \Rightarrow NBU$

Att fördelningen är $IFRA$ betyder att (för $x, y > 0$)

$$\frac{1}{x+y}\int_0^{x+y} \lambda(u)du \geq \frac{1}{x}\int_0^x \lambda(u)du$$

Vi antar nu att $y \leq x$ (i annat fall byter vi x mot y i högerledet ovan). Vi får då

$$\begin{aligned}
\int_0^{x+y} \lambda(u)du \geq \frac{x+y}{x}\int_0^x \lambda(u)du = \\
\int_0^x \lambda(u)du + \frac{y}{x}\int_0^x \lambda(u)du \geq (\text{ty } IFRA \text{ och } x \geq y) \\
\int_0^x \lambda(u)du + \frac{y}{y}\int_0^y \lambda(u)du = \int_0^x \lambda(u)du + \int_0^y \lambda(u)du.
\end{aligned}$$

Detta ser man lätt ger $P(T > x+y) \leq P(T > x)P(T > y)$, dvs NBU gäller.

Alltså gäller $IFR \Rightarrow IFRA \Rightarrow NBU$ och analogt $DFR \Rightarrow DFRA \Rightarrow NWU$. Det finns motexempel på att omvändningarna inte gäller.

6 Poisson-process

Vi vill göra en modell för begreppet “händelser som inträffar fullständigt slumpmässigt i tiden”. Intervallet $(0, t]$ delas i $n = t/h$ st intervall av längd h vardera. Tanken är att låta $h \to 0$.

a) I vardera sådant intervall antar vi att sannolikheten för en händelse är $\lambda h + o(h)$, dvs i princip proportionellt mot intervallets längd. $o(h)$ står för en restterm sådan att den är av mindre storleksordning än h.

b) I vardera sådant intervall är sannolikheten för avsaknad av händelse $1 - \lambda h + o(h)$

c) Sannolikheten för 2 eller fler händelser i ett intervall är $o(h)$, dvs kan försummas.

d) Förekomsten av händelser i olika intervall är oberoende av varandra

Låt $X(t)$ =antalet händelser i intervallet $(0, t]$. Vi har typsituation för binomialfördelning och ser att $X(t)$ är $Bin(n, \lambda h + o(h))$ som blir approximativt $Po(n\lambda h) = Po(\lambda t)$. På motsvarande sätt ser vi att antalet händelser på $(s, t]$ blir $Po(\lambda(t-s))$. Av konstruktionen ser man att tillskotten på disjunkta tidsintervall är oberoende av varandra.
Om vi låter T =tid till första händelsen så erhåller vi

$$P(T > t) = P(X(t) = 0) = e^{-\lambda t}$$

dvs tiden till första händelsen är $Exp(\lambda)$. Tiderna mellan händelserna är oberoende $Exp(\lambda)$-fördelade stoskastiska variabler. Om vi låter

$$S_n = T_1 + T_2 + \cdots + T_n = \text{tiden till } n\text{:te händelsen}$$

erhåller vi på samma sätt

$$P(S_n > t) = P(X(t) \leq n - 1) = \sum_{k=0}^{n-1} \frac{(\lambda t)^k}{k!} e^{-\lambda t}.$$

Om detta deriveras med avseende på t "teleskoperar" uttrycket och vi erhåller

$$f_{S_n}(t) = \frac{\lambda^n t^{n-1}}{(n-1)!} e^{-\lambda t}.$$

som kallas för $\Gamma(n, \lambda)$-fördelningen. Den kan alltså tolkas som fördelningen för summan av n oberoende $Exp(\lambda)$-fördelade stokastiska variabler $T_1, T_2, \ldots, T_n$.

7 Allmänt om analys av livslängdsdata

Vi kommer i detta avsnitt att vilja analysera data om livslängder $X_1, X_2, \ldots, X_n$ för n identiska komponenter där vi alltså ser X_i:na som oberoende likafördelade. Ibland har vi fullständiga data, ibland bara "censurerade" data, dvs bara vissa livslängder observeras. T ex kan man tänka sig att försöket avbryts efter en fix tid t_0 (s k Typ I-censurering) eller att försöket avbryts då r fel observerats (s k Typ II-censurering).

Vi skall ta fram den sk TTT-plotten och TTT-transformen men också skatta felintensitet (om vi tror på att denna är konstant) samt funktionssannolikhet.

Låt $X_1, \ldots, X_n$ vara livslängderna för n komponenter som vi antar är oberoende likafördelade. Låt $X(1), X(2), \ldots, X(n)$ vara dessa ordnade i storleksordning. Alltså är $X(1)$ det minsta av $X_1, X_2, \ldots, X_n$ och $X(2)$ det näst minsta, etc ända till $X(n)$ som är det största. Dessa $X(i)$ brukar kallas ordningsvariablerna.

Vi antar att F har täthet och är strängt växande och att $E(X_i) = \theta$.

Definition: Den total testtiden vid tidpunkt x där $X(i) \leq x < X(i+1)$ är

$$T(x) = \sum_{j=1}^{i} X(j) + (n-i)x.$$

□

Första termen är den testtid som komponenterna som gått sönder före x utsatts för och den sista termen $(n-i)x$ är den testtid som de som klarat sig tiden fram till x utsatts för.

Om vi har numeriska data brukar vi beteckna dem med $x_1, x_2, \ldots, x_n$ som vi ser som utfall av stokastiska variabler $X_1, X_2, \ldots, X_n$ men vi kommer att växla mellan beteckningssätten.

Vi har

$$T(X(i)) = \sum_{j=1}^{i} X(j) + (n-i)X(i)$$

och speciellt $T(X(n)) = \sum_{j=1}^{n} X(j) = \sum_{j=1}^{n} X_n$. Alltså gäller $T(X(n))/n = \bar{X}$. Den relativa testtiden vid i:te felet

$$\frac{T(X(i))}{T(X(n))}$$

och den s k TTT-plotten är en plot av $(i/n, T(X(i))/T(X(n)))$. Vi kommer att visa att

1) Om data kommer från en exponentialfördelning är TTT-plotten ungefär en rät linje

2) Om data kommer från en IFR-fördelning är TTT-plotten ungefär konkav

3) Om data kommer från en DFR-fördelning är TTT-plotten ungefär konvex.

Ordet "ungefär" syftar på att slumpen ger en viss variabilitet i TTT-plotten.

8 Mer egenskaper hos exponentialfördelningen

Om vi vet att $X(a) = 1$ har alltså en händelse inträffat någon gång under intervallet $(0, a]$. Med T =tiden för första händelsen erhålls för $0 \leq t \leq a$

$$P(T \leq t|X(a) = 1) = \frac{P(T \leq t; X(a) = 1)}{P(X(a) = 1)} =$$
$$\frac{P(X(t) = 1; X(a) = 1)}{P(X(a) = 1)} = \frac{P(X(t) = 1; X(a) - X(t) = 0)}{P(X(a) = 1)} =$$
$$\text{(oberoende för disjunkta intervall)} = \frac{P(X(t) = 1)P(X(a) - X(t) = 1)}{P(X(a) = 1)} =$$
$$\frac{\lambda t e^{-\lambda t} e^{-\lambda(a-t)}}{\lambda a e^{-\lambda a}} = \frac{t}{a},$$

som betyder att $T|X(a) = 1$ är likformigt fördelad $R(0, a)$. På samma sätt kan man visa att om det inträffat n händelser på intervallet $(0, a]$ har dessa tider samma fördelning som ordningsvariablerna för n st oberoende likformigt fördelade variabler som är $R(0, a)$. Tiderna inträffar alltså helt slumpmässigt i intervallet.

Sats: Om X_i är $Exp(\lambda_i)$ för $i = 1, 2, \ldots, n$ och oberoende så är $T = \min(X_1, X_2, \ldots, X_n)$ exponentialfördelad $Exp(\sum_1^n \lambda_i)$. $\square$

Bevis:

$$P(T > t) = P(\min(X_1, X_2, \ldots, X_n) > t) = P(X_1 > t; X_2 > t; \ldots; X_n > t) =$$
$$\text{(oberoendet)} = P(X_1 > t)P(X_2 > t) \ldots P(X_n > t) =$$
$$\exp(-\lambda_1 t) \exp(-\lambda_2 t) \ldots \exp(-\lambda_n t) = \exp(-t \sum_{i=1}^{n} \lambda_i),$$

som visar att T är $Exp(\sum_1^n \lambda_i)$. $\square$

Intensiteterna i exponentialfördelningarna adderar sig alltså vid minimumbildning av oberoende exponentialfördelningar.

Man kan också visa (finns som godis) att om L =numret på den som blir minst så gäller att $P(L = j) = \lambda_j / \sum_1^n \lambda_i$ och att L (förvånande nog) är oberoende av T.

Sats: Om $X_1, X_2, \ldots, X_n$ är oberoende $Exp(\lambda)$ så är $T(X(i))/T(X(n))$ för $i = 1, 2, \ldots, n-1$ fördelade som ordningsvariablerna $U(1), U(2), \ldots, U(n-1)$ av $n-1$ oberoende $R(0, 1)$-fördelade variabler $U_1, U_2, \ldots, U_{n-1}$. $\square$

Bevisskiss: Notera att $X(1)$ är minimum av n stycken oberoende $Exp(\lambda)$-fördelade variabler och alltså är $Exp(n\lambda)$. Vidare är $T(X(1)) = nX(1)$ så $T(X(1)) - 0$ är $Exp(\lambda)$. På samma sätt inses att $X(2) - X(1)$ är minmum av $n-1$ oberoende $Exp(\lambda)$ och att alltså $T(X(2)) - T(X(1)) = (n-1)(X(2) - X(1))$ är $Exp(\lambda)$ osv. Vi inser att

$$T(X(1)), T(X(2)) - T(X(1)) \ldots, T(X(n)) - T(X(n-1)) \text{ osv}$$

är oberoende $Exp(\lambda)$ och att alltså $T((X(j))$ är fördelad som summan av j oberoende $Exp(\lambda)$ och alltså utgör tiderna för successiva händelser i en Poissonprocess. Om vi fixerar att $T(X(n)) = a$ kan vi utnyttja resultatet ovan och se att resultatet följer eftersom vi skall placera de $n-1$ händelserna i intervallet $(0, a]$. $\square$

Ur detta erhålls följande resultat:

$$E(\frac{T(X(i))}{T(X(n))}) = \frac{i}{n}$$

och kan visa att spridningen blir ganska liten. Detta betyder att

$$T(X(i))/T(X(n)) \approx i/n.$$

För exponentialfördelning bör TTT-plotten alltså approximativt ge en rät linje. Detta kan användas för att grafiskt undersöka om data kommer från en exponentialfördelning.

9 Empirisk fördelning

Med data $x_1, x_2, \ldots, x_n$ som är utfall av oberoende likafördelade variabler $X_1, X_2, \ldots, X_n$ lägger vi massan $1/n$ i vardera av punkterna $x_1, \ldots, x_n$. Skulle 2 st x_i vara lika hamnar massan $2/n$ i denna punkt, etc. Denna fördelning kallas den empiriska fördelningen för data och vi betecknar dess fördelningsfunktion med $F_n(x)$. Om n är stort liknar den den sanna fördelningen F för X_i:na, dvs att $F_n(x) \approx F(x)$. Formellt blir

$$F_n(x) = \begin{cases} 0 \text{ om } x < x(1) \\ i/n \text{ om } x(i) \leq x < x(i+1) \\ 1 \text{ om } x > x(n) \end{cases}$$

Vi har då följande sats som knyter ihop TTT-plotten med empiriska fördelningen.

Sats: Det gäller att

$$T(x(i)) = n \int_0^{x(i)} (1 - F_n(u))du.$$

Bevis: Ur en figur inses att arean under $F_n(x)$ fram till $x(i)$ är

$$\frac{1}{n} \sum_{k=1}^{i} (x(i) - x(k)) = \frac{i}{n} x(i) - \frac{1}{n} \sum_{k=1}^{i} x(k),$$

där vi beräknar arean som summan av horisontella remsor. Arean under $1 - F_n(x)$ fram till $x(i)$ blir alltså

$$x(i) \cdot 1 - \text{Arean under } F_n \text{ fram till } x(i)$$

och detta ger

$$n\int_0^{x(i)}(1-F_n(u))du = nx(i) - ix(i) + \sum_{j=1}^{i} x(j) = T(x(i)).$$

Om man nu tänker sig att $n \to \infty$ så att F_n närmar sig F och tar $i/n = v$ med $0 \le v \le 1$ så gäller att $x(i) \to F^{-1}(v)$, dvs v-fraktilen i fördelningen. Vi får då

$$\frac{T(x(i))}{n} \to \int_0^{F^{-1}(v)}(1-F(u))du.$$

Speciellt med $v = 1$ dvs $i = n$ erhålls $F^{-1}(v) = \infty$ och VL är alltså $\bar{x}$ medan HL är $E(X)$ så detta svarar precis mot stora talens lag.

Definition: Med TTT-transformen för en fördelning F avses

$$H_F^{-1}(v) = \int_0^{F^{-1}(v)}(1-F(u))du.$$

Speciellt är $H_F^{-1}(1)$ väntevärdet i fördelningen F. $\square$

Den skalade TTT-transformen är $H_F^{-1}(v)/H_F^{-1}(1)$.

Exempel: Om $F(x) = 1 - e^{-\lambda x}$, dvs $Exp(\lambda)$-fördelningen så erhålls inversa funktionen F^{-1} genom att i $y = F(x)$ lösa ut x. Vi får $x = -\frac{1}{\lambda}\ln(1-y)$ så $F^{-1}(v) = -\frac{1}{\lambda}\ln(1-v)$. Vi får alltså

$$H_F^{-1}(v) = \int_0^{-\frac{1}{\lambda}\ln(1-v)} e^{-\lambda u}du = v/\lambda.$$

Eftersom $H_F^{-1}(1) = 1/\lambda$ så erhålls $H_F^{-1}(v)/H_F^{-1}(1) = v$ för $0 \le v \le 1$. $\square$

Sats: Om F är kontinuerlig växande gäller att

$$\left.\frac{d}{dv}H_F^{-1}(v)\right|_{v=F(x)} = \frac{1}{\lambda(x)}$$

där $\lambda(x)$ är felintensiteten, dvs $\lambda(x) = F'(x)/(1-F(x))$. $\square$

Bevis: Om vi låter $G(y) = \int_0^y (1-F(u))du$ är ju $G'(y) = 1 - F(y)$ så enligt kedjeregeln får vi

$$\frac{d}{dv}G(F^{-1}(v)) = G'(F^{-1}(v))\frac{d}{dv}F^{-1}(v),$$

men $u = F^{-1}(v)$, dvs $v = F(u)$ ger $dv = F'(u)du$ och alltså

$$\frac{d}{dv}F^{-1}(v) = \frac{du}{dv} = 1/F'(u) = 1/F'(F^{-1}(v)).$$

Sammantaget ger detta (med $f = F'$)

$$\frac{d}{dv}H_F^{-1}(v) = (1-F(F^{-1}(v)))\frac{1}{f(F^{-1}(v))} = \frac{1-v}{f(F^{-1}(v))}.$$

Sätter vi nu $v = F(x)$ erhålls

$$\frac{d}{dv} H_F^{-1}(v)\bigg|_{v=F(x)} = \frac{1 - F(x)}{f(x)} = \frac{1}{\lambda(x)}.$$

$\square$

Alltså följer följande sats:

Sats:
F är IFR om och endast om $H_F^{-1}(v)$ är konkav.
F är DFR om och endast om $H_F^{-1}(v)$ är konvex. $\square$

Bevis: F är $IFR \Leftrightarrow \lambda(x)$ växer i $x \Leftrightarrow 1/\lambda(x)$ avtar i $x \Leftrightarrow \frac{d}{dv} H_F^{-1}(v)|_{v=F(x)}$ avtar i $x \Leftrightarrow \frac{d}{dv} H_F^{-1}(v)$ avtar i $v \Leftrightarrow H_F^{-1}(v)$ är konkav. $\square$

Om vi nu skattar H_F^{-1} ur data med hjälp av empiriska funktionen och låter

$$H_n^{-1}(v) = \int_0^{F_n^{-1}(v)} (1 - F_n(u))du,$$

ser vi att eftersom $F_n^{-1}(i/n) = x(i)$ och att, som tidigare visats, $T(x(i)) = n \int_0^{x(i)} (1 - F_n(x))dx$ så erhåller vi

$$\frac{H_n^{-1}(i/n)}{H_n^{-1}(1)} = \frac{T(x(i))}{T(x(n))}$$

och vi ser att TTT-plotten är konkav om fördelningen är IFR och konvex om fördelningen är DFR. Man kan alltså använda empiriska TTT-plotten för att avgöra om fördelningen är IFR, DFR eller ingendera. Har den S-form är den varken IFR eller DFR. Badkarskurvan svarar mot att TTT-plotten först är konvex, sedan linjär och till slut konkav. Notera dock att slumpmässigheten i data kan störa detta snygga mönster.

10 Skattning av felintensitet

Vi antar nu att livslängderna är exponentialfördelade $Exp(\lambda)$ och vi vill skatta felintensiteten λ samt göra konfidensintervall för denna parameter.

Vi har täthetsfunktionen $f(x)\lambda e^{-\lambda x}$, $x > 0$. Vi får då vi observerat data $x_1, x_2, \ldots, x_n$ likelihoodfunktionen

$$L(\lambda) = f(x_1)f(x_2)\ldots f(x_n) = \lambda^n \exp(-\lambda \sum_{i=1}^{n} x_i)$$

och loggas detta erhålls

$$\ln L(\lambda) = n \ln(\lambda) - \lambda \sum_{i=1}^{n} x_i$$

som ger

$$\frac{d}{d\lambda}\ln L(\lambda) = \frac{n}{\lambda} - \sum_{i=1}^{n} x_i,$$

så maximum ges av $\lambda = n/\sum_{i=1}^{n} x_j$. Alltså är ML-skattningen av λ

$$\lambda^* = \frac{1}{\bar{X}} = \frac{n}{\sum_{i=1}^{n} X_i} = \frac{n}{T(X(n))}.$$

Vi vet att $T(X(n)) = \sum_1^n X_j$ är $\Gamma(n, \lambda)$, dvs har tätheten

$$f_{T(X(n))}(t) = \frac{\lambda^n t^{n-1}}{(n-1)!}\exp(-\lambda t).$$

Det kan vara lämpligt att införa den s k Γ-funktionen

$$\Gamma(x) = \int_0^\infty t^{x-1}e^{-t}dt$$

där man lätt visar att för positiva heltal n gäller att $\Gamma(n) = (n-1)!$. Vi får alltså med $Y = \sum_1^n X_i$

$$E(\lambda^*) = E(\frac{n}{\sum_{i=1}^{n} X_i}) = nE(\frac{1}{Y}) = n\int_0^\infty \frac{\lambda^n y^{n-1}e^{-\lambda y}}{y\Gamma(n)}dy =$$
$$\frac{n\lambda}{\Gamma(n)}\int_0^\infty (\lambda y)^{n-2}e^{-\lambda y}\lambda dy = (\text{variabelbyte } \lambda y = t) =$$
$$\frac{n\lambda}{\Gamma(n)}\int_0^\infty t^{n-2}e^{-t}dt = \frac{(n-2)!n\lambda}{(n-1)!} = \frac{n}{n-1}\lambda.$$

Alltså är λ^* ej väntevärdesriktig men asymptotiskt väntevärdesriktig. Skattningen

$$\widehat{\lambda} = \frac{n-1}{T(X(n))}$$

är alltså väntevärdesriktig. Med samma metodik kan man visa att

$$V(\widehat{\lambda}) = \frac{\lambda^2}{n-2}.$$

11 Konfidensintervall vid fullständiga data

Sats:

a) X är $Exp(\lambda) \Leftrightarrow \lambda X$ är $Exp(1)$.

b) Y är $\Gamma(n, 1) \Leftrightarrow 2Y$ är $\chi^2(2n)$. □

Bevis: a) $P(\lambda X > x) = P(X > x/\lambda) = \exp(-\lambda x/\lambda) = \exp(-x)$.

b) Enligt Bloms bok har $\chi^2(f)$-fördelningen tätheten

$$\frac{x^{f/2-1}e^{-x/2}}{\Gamma(f/2)2^{f/2}}, \text{ för } x > 0,$$

vilket lätt ger resultatet eftersom $P(2Y \leq x) = P(Y \leq x/2)$ som deriverat ger

$$f_{2Y}(x) = f_Y(x/2)/2 = \frac{(x/2)^{n-1}}{2\Gamma(n)} \exp(-x/2).$$

□

Detta innebär att vi ej behöver ha en tabell över percentiler för Γ-fördelningar utan klarar oss med de sedvanliga tabellerna för χ^2-fördelningarna. I Matlab finns dock Γ-fördelningens percentiler i funktionen gaminv, men denna ingår bara i Stats-modulen som man kanske inte har tillgång till.

Om nu $X_1, X_2, \ldots, X_n$ är oberoende $Exp(\lambda)$ gäller alltså att

$$2\lambda \sum_{i=1}^{n} X_i \text{ är } \chi^2(2n).$$

Vi får alltså

$$1 - \alpha = P\left(\chi^2_{1-\alpha/2}(2n) \leq 2\lambda \sum_{i=1}^{n} X_i \leq \chi^2_{\alpha/2}(2n)\right),$$

som ger ett konfidensintervall för λ med konfidensgrad $1 - \alpha$ till

$$\left(\frac{\chi^2_{1-\alpha/2}(2n)}{2\sum_{i=1}^{n} x_i}, \frac{\chi^2_{\alpha/2}(2n)}{2\sum_{i=1}^{n} x_i}\right) = \left(\frac{\chi^2_{1-\alpha/2}(2n)}{2T(x(n))}, \frac{\chi^2_{\alpha/2}(2n)}{2T(x(n))}\right).$$

Notera att om man vill ha konfidensintervall för $m = 1/\lambda$ är det bara att ta 1/gränserna ovan. Allmänt gäller att om man har ett konfidensintervall för en parameter θ och vill ha ett konfidensintervall för $\psi = g(\theta)$ där g är en monoton (växande eller avtagande funktion) tar man bara g av gränserna i intervallet för θ. Eventuellt kan gränserna, som för $1/\lambda$ ovan, kastas om.

12 Censurerade data

Vi har 4 olika typer av censurering, dvs situationer då vi inte har fullständiga data.

Typ I-censurering: Försöket avslutas vid en fix tid t_0.

Typ II-censurering: Försöket avslutas då r av de n enheterna felat, dvs vid tiden $X(r)$.

Typ III-censurering: Försöket avslutas vid $\min(t_0, X(r))$, dvs vid tiden t_0 eller (om detta inträffar tidigare) då r felat.

Typ IV-censurering: Oberoende stokastisk censurering av enheterna, dvs varje observation censureras slumpmässigt. För en censurerad observation vet vi alltså bara att livslängden är större än censureringstiden.

Vi kommer i detta avsnitt att behandla censurering av Typ I och Typ II. Vi kommer också att för dessa studera två olika fall:

a) Med återläggning, där felande komponenter ersätts med nya

b) Utan återläggning, där felande enheter ej ersätts.

Vi är särskilt intresserade av total testtid, TTT, eftersom denna kan användas för skattningsändamål.

Typ IV-censurering behandlas i nästa avsnitt som behandlar skattning av överlevnadsfunktion och felintensitet.

12.1 Typ II censurering utan återläggning

Försöket avbryts då r enheter felat. Vi observerar alltså $x(1), x(2), \ldots, x(r)$ där r är fixt. Resten av enheterna är censurerade – vi vet bara att de felar efter $X(r)$ då försöket avslutas. Vi har alltså r st fel av n testade och felande ersätts ej. Total testtid blir

$$T = \sum_{j=1}^{r} x(j) + (n-r)x(r).$$

Vi observerar de r minsta av $x_1, \ldots, x_n$ och för de övriga $n-r$ bara observerar att de är större än den r:te i storleksordning, dvs $x(r)$. Likelihood-funktionen består då dels av produkten av tätheten i $x(1), x(2), \ldots, x(r)$ och dels sannolikheten att de övriga $n-r$ är större än $x(r)$. Denna sista sannolikhet är $P(X_i > x(r)) = \exp(-\lambda x(r))$ och vi får alltså likelihoodfunktionen

$$L(\lambda) = \lambda^r \exp(-\lambda \sum_{j=1}^{r} x(j)) \exp(-\lambda(n-r)x(r)) = \lambda^r \exp(-\lambda T(x(r)))$$

där (som tidigare) $T(x(r))$ är totala testtiden fram till r:te felet. Vi får lätt att ML-skattningen är

$$\lambda^* = \frac{r}{T(x(r))},$$

där $T(X(r)) = \sum_{j=1}^{r} X(j) + (n-r)X(r)$. Tiden från 0 till $X(1)$ är minimum av n st oberoende $Exp(\lambda)$ är alltså $Exp(n\lambda)$. På samma sätt är $X(2) - X(1)$ minimum av $n-1$ oberoende $Exp(\lambda)$ och alltså $Exp((n-1)\lambda)$. Allmänt är $D_j = X(j) - X(j-1)$ minimum av $n-j+1$ oberoende $Exp(\lambda)$ och är alltså $Exp((n-j+1)\lambda)$ för $j = 1, 2, \ldots, r$. Men

$$T(X(r)) = nD_1 + (n-1)D_2 + \cdots + (n-r+1)D_r$$

och detta ger att $\lambda T(X(r))$ är $\Gamma(r, 1)$. Vi får alltså (som vid fullständiga data) att $2\lambda T(X(r))$ är $\chi^2(2r)$. Vi inser också att λ^* ej är väntevärdesriktig men att $E(\lambda^*) = r\lambda/(r-1)$. Vi får samma typ av konfidensintervall för λ som för fullständiga data med mindre modifikationer

$$\left(\frac{\chi^2_{1-\alpha/2}(2r)}{2T(x(r))}, \frac{\chi^2_{\alpha/2}(2r)}{2T(x(r))} \right).$$

12.2 Typ II-censurering med återläggning

Vi har total testtid $T(x(r)) = nx(r)$ eftersom alla n provbänkar är besatta hela tiden fram till vi avbryter försöket då r fel uppstått. Tiderna mellan felen ($D_j = X(j) - X(j-1)$ för $j = 1, 2, \ldots, r$) är minimum av n oberoende $Exp(\lambda)$ och alltså $Exp(n\lambda)$ och den totala testtiden kan skrivas

$$T(X(j)) = n(D_1 + D_2 + \cdots + D_r).$$

Återigen gäller alltså att $T(X(r))$ är $\Gamma(r, \lambda)$ och därmed att $2\lambda T(X(r))$ är $\chi^2(2r)$. Vi erhåller samma konfidensintervall som i föregående fall.

12.3 Typ I censurering med återläggning

Så snart en enhet felat ersätts den med en ny. Alltså är hela tiden n under provning. Total testtid blir $T = nt_0$. Antalet fel är s som alltså är slumpmässigt.

Vi får ML-skattningen $\lambda^* = s/(nt_0)$ där s är antalet observerade fel. Notera att s är slumpmässigt och ett utfall av S som är $Po(\lambda nt_0)$. Alltså är λ^* väntevärdesriktig eftersom $E(S) = \lambda nt_0$. Vidare är $V(S) = \lambda nt_0$ som skattas med $\lambda^* nt_0 = s$ och vi får medelfelet $d(s) = \sqrt{s}$.

Vi skulle kunna göra ett approximativt konfidensintervall genom normalapproximation av Poisson-fördelningen och erhålla det approximativa intervallet $s \pm \lambda_{\alpha/2}\sqrt{s}$ för λnt_0 dvs intervallet

$$\frac{s}{nt_0} \pm \frac{\lambda_{\alpha/2}\sqrt{s}}{nt_0}$$

för λ åtminstone om s är stort, t ex $s \geq 15$.

Oftast ignorerar man dock att censureringen är av Typ I och anger konfidensintervallet enligt Typ II-censurering dvs som

$$\left(\frac{\chi^2_{1-\alpha/2}(2s)}{2T(x(s))}, \frac{\chi^2_{\alpha/2}(2s)}{2T(x(s))} \right),$$

som alltså har approximativ konfidensgrad $1 - \alpha$.

12.4 Typ I-censurering utan återläggning

Försöket avbryts vid tid t_0. Av n enheter som provas går s (slumpmässigt antal) sönder. Vi observerar $X(1), X(2), \ldots, X(s)$.

Vi har totala testtiden

$$T(X(S)) = \sum_{j=1}^{S} X(j) + (n - S)t_0$$

där S är det stokastiska antal fel som inträffar före t_0. Skattningen $\lambda^* = s/T(x(s))$ är en approximativt väntevärdesriktig skattning. Att beräkna ett exakt konfidensintervall är besvärligt, men man brukar då i stället göra som för Typ I-censurering med återläggning, dvs bortse från att antalet fel är slumpmässigt och återigen ta intervallet

$$\left(\frac{\chi^2_{1-\alpha/2}(2s)}{2T(x(s))}, \frac{\chi^2_{\alpha/2}(2s)}{2T(x(s))}\right).$$

13 Skattning av överlevnadsfunktion

13.1 Fullständigas data

Med fullständiga data $x_1, x_2, \ldots, x_n$ skattar man naturligtvis den bakomliggande fördelningen med den empiriska fördelningen $F_n(x)$, dvs med den fördelning som lägger massan $1/n$ i vardera observation. Överlevnadsfunktionen skattas med $\widehat{R}(x) = 1 - F_n(x)$.

Vi övergår nu till data som har censurering av Typ IV och anger två olika skattningar, Kaplan-Meier-skattningen (ibland kallad ”product limit estimator”) och Nelson-estimatorn.

13.2 Kaplan-Meier-estimatorn

Vi antar att censureringen sker vid stokastiska tidpunkter, dvs att varje enskilt data x_i antingen anger tidpunkten för fel eller anger en tidpunkt fram till vilken komponenten fungerat. Vi har alltså ordnade observationer $x(1) \leq x(2) \leq \cdots \leq x(n)$ där vissa av dessa är censurerade. Låt ν genomlöpa de värden j där $x(j)$ är en tid till fel och $x(j) < x$. Kaplan-Meier-estimatorn av överlevnadsfunktionen är då

$$\widehat{R}(x) = \prod_{\nu} \frac{n-\nu}{n-\nu+1}.$$

För att motivera detta gör vi följande betraktelse. Vi tänker oss att tiden mellan 0 och x delas in i ett stort antal korta delintervall $0 = u_0, u_1, u_2, \ldots, u_m = x$. Vi antar att indelningen är så fin att bara en av de enskilda händelserna fel respektive censurering inträffar i de enskilda intervallen. Vi har

$$\begin{aligned} R(x) = P(X > x) = P(X > u_m) = \\ P(X > u_m | X > u_{m-1}) P(X > u_{m-1} | X > u_{m-2}) \ldots \\ \ldots P(X > u_2 | X > u_1) P(X > u_1 | X > u_0). \end{aligned}$$

Vi vill nu skatta de enskilda faktorerna i denna produkt.

a) I ett intervall där varken fel eller censurering inträffat skattas den med 1.

b) Om censurering men inte fel inträffat skattas den också med 1.

c) Om ett fel inträffat och man haft $n-r+1$ enheter igång vid början av intervallet och $n-r$ i slutet skattas den med $(n-r)/(n-r+1)$.

Vi får då

$$\widehat{R}(x)=1\cdot 1\cdot 1\cdot \frac{n-\nu_1}{n-\nu_1+1}\cdot 1\cdots 1\cdot \frac{n-\nu_2}{n-\nu_2+1}\cdots=\prod_{\nu}\frac{n-\nu}{n-\nu+1},$$

där $x(\nu_1), x(\nu_2), \ldots$ är de tidpunkter som komponenter felar. Om data skulle sakna censureringar överensstämmer denna estimator med $1-F_n(x)$. Med censureringar hackar överlevnadskurvan vid varje observerat fel ner med en faktor som är andelen som icke-felande av de kvarvarande ocensurerade som ”lever” vid den tidpunkten.

13.3 Nelson-estimatorn

Vi låter $\Lambda(x)=\int_0^x \lambda(u)du$ och noterar att $R(x)=e^{-\Lambda(x)}$ och alltså

$$\Lambda(x)=-\ln(R(x))=-\ln(1-F(x)).$$

Vi försöker nu skatta $\Lambda(x)$ ur data. Vi börjar med en enkel sats.

Sats: Låt X vara kontinuerlig med strängt växande fördelningsfunktion F. Då gäller

a) $U=F(X)$ är rektangelfördelad $R(0,1)$.

b) $Z=-\ln(1-F(X))$ är $Exp(1)$-fördelad. $\Box$

Bevis: a) Med $0<x<1$ erhålls

$$P(U\le x)=P(F(X)\le x)=P(X\le F^{-1}(x))=F(F^{-1}(x))=x,$$

som ju är fördelningsfunktion för $R(0,1)$-fördelningen.

b) Med $z>0$ erhålls

$$P(Z\le z)=P(-\ln(1-F(X))\le z)=P(1-F(X)\ge e^{-z})=\\ P(F(X)\le 1-e^{-z})=P(U\le 1-e^{-z})=1-e^{-z},$$

som innebär att Z är $Exp(1)$. $\Box$

Notera att transformationen som ger Z är samma som dök upp ovan i uttrycket för $\Lambda(x)$. Detta innebär att vi på sätt och vis transformerar data så att de svarar mot ordningsvariabler för exponentialfördelningar.

Vi inleder med situationen med fullständiga data, dvs utan censurering. Låt $X_1, X_2, \ldots, X_n$ vara oberoende likafördelade med fördelning enligt F. Motsvarande ordningsvariabler kallar vi som tidigare $X(1), X(2), \ldots, X(n)$. Vi bildar nu $Z_j=-\ln(1-F(X_j))$ enligt satsen ovan och noterar att alltså Z_j;na är oberoende $Exp(1)$. Låt på motsvarande sätt $Z(j)=-\ln(1-F(X(j)))$ vara ordningsvariablerna för Z:variablerna. De utgör alltså ordningsvariabler för n oberoende $Exp(1)$-variabler. Vi har

$$E(Z(j)) = \frac{1}{n} + \frac{1}{n-1} + \cdots + \frac{1}{n-j+1}.$$

Detta eftersom $Z(1)$ är minimum av n oberoende $Exp(1)$ och alltså är $Exp(n)$, och vidare $Z(2) - Z(1)$ är minimum av $n-1$ oberoende $Exp(1)$ och alltså är $Exp(n-1)$ osv.

Notera att $\Lambda(X(j)) = -\ln(1 - F(X(j))) = Z(j)$ så att alltså $E(\Lambda(X(j))) = E(Z(j))$. Vi skattar då $\Lambda(x)$ med

$$\widehat{\Lambda}(x) = \sum_{i=1}^{j} \frac{1}{n-i+1} \text{ då } X(j) \leq x < X(j+1)$$

som ger ungefär "rätt" väntevärde, dvs $E(\widehat{\Lambda}(X(j))) = E(\Lambda(X(j)))$. Vi skattar sen $R(x)$ med $\widehat{R}(x) = e^{-\widehat{\Lambda}(x)}$.

Skulle vi ha censurerade data modifierar vi detta så att

$$\widehat{\Lambda}(x) = \sum_{\nu} \frac{1}{n-\nu+1}$$

då ν genomlöper de j där $x(j)$ är tid för ett fel och $x(j) < x$. Rent tekniskt ökar $\widehat{\Lambda}$ vid observerade fel med en term som är 1/(antalet i drift vid tidpunkten). Överlevnadsfunktionen skattas sen återigen med $\widehat{R}(x) = e^{-\widehat{\Lambda}(x)}$. Som för Kaplan-Meier-estimatorn hackar $\widehat{R}$ ner med en faktor vid varje observerat fel. Denna faktor är $\exp(-1/(n-\nu+1))$ fär $n-\nu+1$ är antalet i drift vid den aktuella tidpunkten. En enkel serieutveckling

$$e^{-1/(n-\nu+1)} \approx 1 - \frac{1}{n-\nu+1} = \frac{n-\nu}{n-\nu+1}$$

visar att Kaplan-Meier-skattningen och Nelson-skattningen är rätt lika särskilt om antalet fel och censureringar är litet i förhållande till n.