7.) a) Za (ne nujno zvezno) s. s. X s k. p. f. F čim bolj natančno obravnavajte k. p. f. slučajne spremenljivke $Y = F(X)$.

• Če je X zvezna s. s. sledi, da je tudi $F = F_X$ zvezna funkcija: $\mathbb{R} \to [0,1]$.
Tedaj je $Y = F(X) \sim U(0,1)$.

Vzemimo $y \in (0,1)$. Tedaj je $F^{-1}(y) = [X', X]$ (to je interval zaradi monotonosti, zaprt zaradi zveznosti, omejen zaradi $\lim_{x\to\infty} F(x) = 1$ in $\lim_{x\to-\infty} F(x) = 0$)

Zato sledi $F^{-1}((-\infty, y]) = F^{-1}([0,y]) = (-\infty, x]$.

Sledi $P(F(X) \leq y) = P(F(X) \in (-\infty, y]) = P(X \in F^{-1}((-\infty,y])) = P(X \in (-\infty, x]) =$
$= P(X \leq x) = F(x) = y$

Po drugi strani: ~~ke~~ ker je zvezna, vemo, da ~~...~~ invert $F_X^{-1}(y)$ obstaja. Sledi
$P(Y \leq y) = P(F_X(X) \leq y) = P(X \leq F_X^{-1}(y)) \underset{\downarrow}{=} F_X(F_X^{-1}(y)) = y$
F_X je naraščajoča

$\hookrightarrow$ Torej je za $\forall y \in [0,1]$ $Y = F_X(X) \sim U(0,1)$.

• Če X ni zvezna s. s.:
Tudi F_X ~~...~~ ni nujno zvezna, vendar ker gre za kumulativno porazdelitveno funkcijo, mora zanjo veljati:

1.) $\lim_{x\to\infty} F_X(x) = 1$ 2.) $\lim_{x\to-\infty} F_X(x) = 0$ 3.) $F_X(x)$ je (ne nujno strogo) naraščajoča
$\downarrow$
~~to~~ to je glavna pomanjkljivost konstrukcije za inverza.

4.) F_X je desno zvezna

Za $Y \sim F_X(x)$ velja $F_Y(y) = P(Y \leq y) = P(F_X(X) \leq y)$
Vemo: $\forall z \in \mathbb{R}: F_X(z) \in [0,1]$

Vrednosti F_Y na $(-\infty, 0]$ in $[1,\infty)$ lahko hitro določimo: $F_Y(y) = \begin{cases} 0 ; & y \in (-\infty, 0] \\ 1 ; & y \in [1,\infty) \\ * ; & y \in (0,1) \end{cases}$

Moramo določiti (le še) $*$, torej imamo $y \in (0,1)$.

Ideja: $F_X(x) \leq y \iff X \leq F_X^{-1}(y)$, kjer pa F_X^{-1} ne obstaja nujno.

Spomnimo se na splošni invers kumulativne porazdelitvene funkcije, imenovan tudi funkcija kvantilov, definirana kot: $F^{-1}(p) = \inf\{x \mid F_X(x) \geq p\}$ $\forall p \in [0,1]$.

Lastnosti takšnega inverza so:

• F^{-1} je nepadajoča
• $F^{-1}(F(x)) \leq x$
• $F(F^{-1}(p)) \geq p$
• $F^{-1}(p) \leq x \iff p \leq F(x)$

• Če je $\{X_\alpha\}$ nabor neodvisnih s. s. porazdeljenih s F, potem ~~obstajajo...~~ obstajajo takšne s. s. Y_α, da je $Y_\alpha \sim U[0,1]$ in velja $F^{-1}(Y_\alpha) = X_\alpha$ z verjetnostjo 1

To je ugodno, saj ni odvisno od zveznosti X.

Torej, če za naš primer definiramo $G(y) = \inf\{x \mid F_X(x) \geq y\}$ $\quad y \in (0,1)$
je to dobro definirano, saj tak x obstaja, saj veljajo lastnosti kumulativne porazdelitvene funkcije, še več, po teh lastnostih vrednost x celo zavzamemo (zaradi naraščajočosti). $\Rightarrow \inf \to \min$

Poračunajmo:

$$F_Y(y) = P(F_X(X) \leq y) = P(X \leq G(y)) = F_X(G(y)) = F_X(\min\{x \mid F_X(x) \geq y\}) =$$

monotonost

$$= \min(F_X(\{x \mid F_X(x) \geq y\})) =$$

$$= \min\{F_X(x) \mid F_X(x) \geq y\} \overset{\geq}{\cancel{=}} y \;//$$

To ni nujno res, saj $F_X(x)$ ne zavzame nujno vseh y. Za tiste, ki jih zavzame, ležijo na simetrali lihih kvadrantov na $[0,1]$, ostali ležijo nad njo.

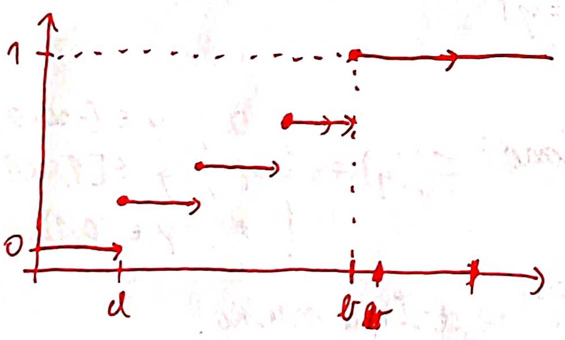
~~$\Rightarrow$ Pričakovano sledi~~

$$\cancel{F_Y(y) = \begin{cases} 0 \; ; & y \in (-\infty, 0] \\ 1 \; ; & y \in [1, \infty) \\ y \; ; & y \in (0,1) \end{cases}} \sim U[0,1].$$

Pogoj k zveznosti bi nam ravno zagotovil, da bi zajeli vse y.
Ali bi lahko to vseeno posplošili z diskretno enakomerno porazdelitvijo:

$$P(X=k) = \begin{cases} \frac{1}{n} \; ; & k = x_i, \; i \in \{1,2,\dots,n\} \\ 0 \; ; & \text{sicer} \end{cases}$$

Graf zbirne funkcije:



Zdi se, kot da bi lahko delovalo.

7. b) Dokažite (i) točko splošnega izreka o konstrukciji I. Z. v primeru enoparametričnega modela.

Izrek: Naj bo $\vartheta \in \Theta \subset \mathbb{R}$ realnoštevilski parameter modela za X z vrednostmi v $\mathbb{R}^n$. Naj bo $T: \mathbb{R}^n \to \mathbb{R}$ statistika (npr. zadostna statistika v enoparametričnem eksponentnem modelu) in naj bodo $F_{T;\vartheta}$ pridružene k. p. f. Razcepimo $\alpha = \alpha_1 + \alpha_2$.

(i) Naj bosta $F_{T;\vartheta}(t)$ in $F_{T;\vartheta}(t^-)$ nenaraščajoči v ϑ za $\forall$ fiksen t. Definiramo:

$\bar{\vartheta} = \sup\{\vartheta \mid F_{T;\vartheta}(t) \geq \alpha_1\}$ in $\underline{\vartheta} = \inf\{\vartheta \mid F_{T;\vartheta}(t^-) \leq 1-\alpha_2\}$

Tedaj je $[\underline{\vartheta}(TX), \bar{\vartheta}(TX)]$ interval zaupanja stopnje zaupanja α za ϑ.

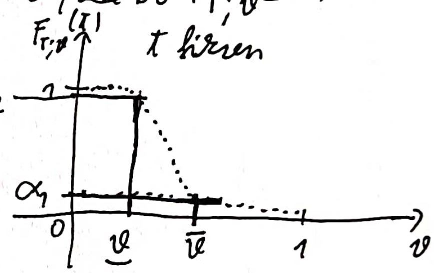
Dokaz: Pokazati moramo, da je interval $[\underline{\vartheta}(TX), \bar{\vartheta}(TX)]$ res interval zaupanja s stopnjo zaupanja α za parameter ϑ.

Po definiciji $\underline{\vartheta}$ in $\bar{\vartheta}$ vemo: $F_{T;\underline{\vartheta}}(t^-) \leq 1-\alpha_2$ in $F_{T;\bar{\vartheta}}(t) \geq \alpha_1$

To pomeni: $F_{T;\vartheta}(t)$ je pridružena k. p. f. za statistiko T. Opazujemo jo pri fiksnem t, ki je realizacija statistike T. ~~Za $\underline{\vartheta}$ iščemo najmanjšo tako ϑ, da bo $F_{T;\vartheta}(t)$~~

za $\bar{\vartheta}$ iščemo največjo tako ϑ, da bo $F_{T;\vartheta}(t) \geq \alpha_1$.

Ker je $F_{T;\vartheta}$ nenaraščajoča, njena vrednost z višjimi vrednostmi ϑ pada.



To pomeni: ~~$P(\vartheta \leq \underline{\vartheta}) \leq 1-\alpha_1$ in $P(\vartheta \geq \bar{\vartheta}) \leq \alpha_2$~~

$$P(\vartheta > \bar{\vartheta}) \leq \alpha_1 \quad \text{in} \quad P(\vartheta < \underline{\vartheta}) \leq \alpha_2 \qquad (*)$$

Zanima nas verjetnost, da $\vartheta(TX)$ pade izven $[\underline{\vartheta}, \bar{\vartheta}]$. To zapišemo kot

$$P(\vartheta(TX) < \underline{\vartheta}) + P(\vartheta(TX) > \bar{\vartheta}) \underset{(*)}{=} P(\vartheta < \underline{\vartheta}) + P(\vartheta > \bar{\vartheta}) \leq \alpha_2 + \alpha_1 = \alpha < 1$$

Zgornja neenačba nam pove, da je verjetnost da $\vartheta(TX)$ pade izven $[\underline{\vartheta}, \bar{\vartheta}]$ manjša ali enaka $\alpha_2 + \alpha_1 = \alpha$. To pomeni, da je verjetnost, da pade v interval $[\underline{\vartheta}, \bar{\vartheta}]$:

~~$P(\underline{\vartheta}(TX) < \vartheta(TX) < \vartheta(TX)$~~

$$P(\underline{\vartheta}(TX) < \vartheta(TX) < \bar{\vartheta}(TX)) = 1 - P(\vartheta(T(X)) > \bar{\vartheta}) - P(\vartheta(T(X)) < \underline{\vartheta}) \geq 1 - (\alpha_1 + \alpha_2) = 1 - \alpha.$$

Sledi, da je $[\underline{\vartheta}(T(X)), \bar{\vartheta}(T(X))]$ I. Z. s stopnjo zaupanja α za parameter ϑ.