Leopea Cătălina, grupa 243

Tema 4

Exercițiul 1

a) X v.a.

$$f(x) = \begin{cases} c\ln\left(\frac{7}{x}\right), & 0<x<7 \\ 0, & \text{altfel} \end{cases}$$ densitate de probabilitate $\Leftrightarrow$

$$\Leftrightarrow \int_{-\infty}^{\infty} f(x)\,dx = 1$$

$$\int_{-\infty}^{\infty} f(x)\,dx = \int_0^7 f(x)\,dx = \int_0^7 c\ln\left(\frac{7}{x}\right)dx = \int_0^7 c\ln\left(\frac{7}{x}\right)dx = c\int_0^7 \ln\left(\frac{7}{x}\right)dx$$

$$\int_0^7 \ln\left(\frac{7}{x}\right)dx \overset{\substack{f'=1 \\ g=\ln(7/x)}}{=} x\ln\left(\frac{7}{x}\right) - \int x\cdot\frac{x}{7}\cdot\left(\frac{7}{x}\right)' \Bigg|_0^7 =$$

$$= x\ln\left(\frac{7}{x}\right) - \int x\cdot\frac{x}{7}\cdot 7\left(-\frac{1}{x^2}\right)\Bigg|_0^7 = x\ln\left(\frac{7}{x}\right) + x\Bigg|_0^7 =$$

$$= 7\ln(1) + 7 - \lim_{x\to 0} x\ln\left(\frac{7}{x}\right) = 7 - \lim_{x\to 0} x\ln\left(\frac{7}{x}\right) = 7-0=7$$

$$\lim x\ln\left(\frac{7}{x}\right) = \lim_{x\to 0^+} x\ln 7 - x\ln x = \lim_{x\to 0^+} \frac{-\ln x}{\frac{1}{x}}$$

$$= \lim_{x\to 0^+} -\frac{\frac{1}{x}}{-\frac{1}{x^2}} = \lim_{x\to 0^+} x = 0$$

$$\int_0^7 c\ln\left(\frac{7}{x}\right)dx = 1 \Leftrightarrow 7c = 1 \Leftrightarrow c = \frac{1}{7}$$

$$F(x) = \begin{cases} 0, & x\le 0 \\ \int_0^x f(x)\,dx, & x\in(0,7) \\ 1, & x\ge 7 \end{cases}$$

$$\int_0^x f(x)\,dx = \int_0^x \frac{1}{7}\ln\left(\frac{7}{x}\right)dx = \frac{1}{7}\left(x\ln\left(\frac{7}{x}\right)+x\right) = \frac{x}{7}\left[\ln\left(\frac{7}{x}\right)+1\right]$$

b) $$f(x) = \begin{cases} \frac{1}{x}, & x\in[1-c, 1+c] \\ 0, & \text{altfel} \end{cases}$$

$$\int_{-\infty}^{\infty} f(x) = \int_{1-c}^{1+c} \frac{1}{x}\,dx = 1$$

$$\int_{1-c}^{1+c} \frac{1}{x}\,dx = \ln(1+c) - \ln(1-c) = \ln\left(\frac{1+c}{1-c}\right) = \ln e$$

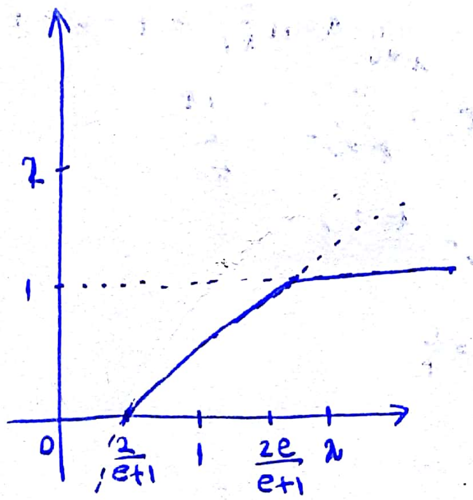
$$\Rightarrow \frac{1+c}{1-c} = e \Rightarrow c = -\frac{1-e}{1+e} = \frac{e-1}{e+1}$$

$$1-c = 1 - \frac{e-1}{e+1} = \frac{2}{e+1}$$

$$1+c = 1 + \frac{e-1}{e+1} = \frac{2e}{e+1}$$

$$\int_{1-c}^{x} f(x) = \ln(x) - \ln(1-c) = \ln(x) - \ln\left(\frac{2}{e+1}\right)$$

$$F(x) = \begin{cases} 0, & x < \frac{2}{e+1} \\ \ln(x) - \ln\left(\frac{2}{e+1}\right), & x \in \left[\frac{2}{e+1}, \frac{2e}{e+1}\right] \\ 1, & x > \frac{2e}{e+1} \end{cases}$$



Exercitiul 2

X v.a. pentru zgomotul unei masini de spălat

$E(X) = 44$ dB $= \mu$

σ = deviatia standard $= 5$ dB

$P(X > 48\text{ dB}) = ?$ într-un esantion de 10 masini

Fie X_i - v.a. pt zgomotul produs de mașina i

$X_1, X_2, \dots, X_{10}$ v.a. independente cu $\mu = E(X_i) = 44$ dB

$\sigma = 5dB$

Fie $\overline{X_n}$ media $X_1, X_2, \dots, X_{10}$ media zgomotului

$$\overline{X_n} = \frac{X_1 + X_2 + \dots + X_{10}}{n}. \text{ Avem } E(\overline{X_n}) = E\left(\frac{X_1 + X_2 + \dots + X_{10}}{n}\right)$$

$$= \frac{1}{n}\left[E(X_1) + E(X_2) + \dots + E(X_{10})\right] = \frac{1}{n} \cdot n \cdot \mu = \mu = 44$$

$Var(X_i) = \sigma^2 = 25$

$$Var(\overline{X_n}) = Var\left(\frac{X_1 + X_2 + \dots + X_n}{n}\right) = \frac{1}{n^2}\left[Var(X_1) + \dots + Var(X_n)\right]$$

$$= \frac{1}{n^2} \cdot n\sigma^2 = \frac{\sigma^2}{n} \Rightarrow \sigma_{\overline{X_n}} = \sqrt{\frac{\sigma^2}{n}} = \frac{\sigma}{\sqrt{n}} = \frac{5}{\sqrt{10}}$$

Cf Teoremei de limită centrală, standardizarea lui $\overline{X_n}$ e aprox normala standard $(N(0,1))$.

$$Z_n = \frac{\overline{X_n} - \mu_{\overline{X_n}}}{\sigma_{\overline{X_n}}} = \text{standardizarea lui } \overline{X_n}$$

$$\Rightarrow P(\overline{X_n} > 48 dB) = P\left(\frac{\overline{X_n} - 44}{\frac{5}{\sqrt{10}}} > \frac{48 - 44}{\frac{5}{\sqrt{10}}}\right)$$

$$\simeq P\left(Z > \frac{4\sqrt{10}}{5}\right) = 0.0057 \text{ (calculat în R cu pnorm)}$$

în $N(0,1)$

b) $n = 100$

X_i = greutatea unei persoane

$E(X_i) = 66.3 = \mu$

$\sigma = 15.6$

$S_{100} = X_1 + X_2 + \dots + X_{100}$

$P(S_{100} > 7000) = ?$

$E(S_{100}) = E(X_1) + \dots + E(X_{100}) = n\mu = 6630$

$Var(S_{100}) = Var(X_1) + \dots + Var(X_{100}) = n \cdot \sigma^2$

$\sigma_{S_{100}} = \sqrt{n\sigma^2} = \sigma\sqrt{n} = 156$

Conform Teoremei de limită centrală

$$P(S_{100} > 7000) = P\left(\frac{S_{100} - \mu_{S_{100}}}{\sigma_{S_{100}}} > \frac{7000 - \mu_{S_{100}}}{\sigma_{S_{100}}}\right)$$

$\simeq P\left(Z > \frac{7000 - 6630}{156}\right)$ în $N(0,1)$

$= P(Z > 2.371) \simeq 0.00885$

Exercițiul 3

$n = ?$ a.î. X_i = o aruncare a.î.

X_i	0	1
$f_{X_i}(X_i)$	0.5	0.5

$E[X_i] = 0.5$

$Var(X_i) = p(1-p) = 0.5 \cdot 0.5 = \frac{1}{4} \Rightarrow \sigma = \frac{1}{2}$

$\overline{X_n} = \frac{X_1 + X_2 + \ldots + X_n}{n}$

Frecvența de apariție a stemei este media lui $X_1 \ldots X_n$

$E[\overline{X_n}] = \frac{1}{n}[E(X_1) + \ldots + E(X_n)] = \frac{1}{n} \cdot n \cdot \frac{1}{2} = \frac{1}{2}$

$V[\overline{X_n}] = \frac{1}{n^2} \cdot n \cdot \frac{1}{4} = \frac{\sigma^2}{n}$

$\sigma_{\overline{X_n}} = \frac{\sigma}{\sqrt{n}} = \frac{1}{2\sqrt{n}}$

$P(0.49 < \overline{X_n} < 0.51) = 0.6$

Conform Teoremei centrale de medie

$\overline{X_n} \simeq N\left(\mu_{\overline{X_n}}, \frac{\sigma^2}{n}\right)$ sau standardizat $\frac{\overline{X_n} - \mu_{\overline{X_n}}}{\frac{\sigma}{\sqrt{n}}} \simeq N(0,1)$

$P(0.49 < \overline{X_n} < 0.51) = P\left(\frac{0.49 - 0.5}{\frac{1}{2\sqrt{n}}} < Z < \frac{0.51 - 0.5}{\frac{1}{2\sqrt{n}}}\right)$ în $N(0,1)$

$P(-0.01 \cdot 2\sqrt{n} < Z < 0.01 \cdot 2\sqrt{n}) = 0.6$

Deci $-0.01 \cdot 2\sqrt{n}$ e quantila de ordin 0.2

$+0.01 \cdot 2\sqrt{n}$ e quantila de ordin 0.8

qnorm $(0.2, 0, 1) \simeq -0.8416$

$-0.8416 = -0.01 \cdot 2 \cdot \sqrt{n} \Rightarrow$

$\Rightarrow \sqrt{n} = \frac{0.8416}{0.02} = 42.018 \Rightarrow n \simeq 1770$

qnorm $(0.8, 0, 1) = 0.8416 \Rightarrow n$ e același

Exercițiul 4

a) $X_1, X_2 \dots X_n$ cu rep $F(x)$
dens $f(x)$

$Y_1, Y_2 \dots Y_n$ versiune ordonată crescător

$H_k(x) = f$ repartiție a lui Y_k

$h_k(x) = f$ de distribuție a lui Y_k

$Y_1 = \inf X$

$H_1(x) = f$ de repartiție a lui Y_1

$H_1(x) = P(Y_1 < x) = 1 - P(Y_1 > x)$

$$P(Y_1 > x) = P((X_1 > x) \cap (X_2 > x) \cap \dots \cap (X_n > x))$$
$$= P(X_1 > x) \cdot P(X_2 > x) \cdot \dots \cdot P(X_n > x)$$
$$= [1 - F(x)]^n$$

$$\Rightarrow H_1(x) = 1 - [1 - F(x)]^n$$

$$h_1(x) = (1 - [1 - F(x)]^n)' = (-1) \cdot n \cdot [1 - F(x)]^{n-1} \cdot (1 - F(n))'$$
$$= n \cdot [1 - F(x)]^{n-1} \cdot f(x)$$

$Y_n = \sup X_i$

$H_n(x) = f$ de distribuție

$$H_n(x) = P(Y_1 < x) = P((X_1 < x) \cap (X_2 < x) \cap \dots \cap (X_n < x))$$
$$= P(X_1 < x) \cdot P(X_2 < x) \cdot \dots \cdot P(X_n < x)$$
$$= F(x)^n$$

$$h_n(x) = (F(x)^n)' = nF(n)^{n-1} f(x)$$

b) $P(Z > \mu + 3\sigma)$ în $N(\mu, \sigma^2)$

$$P\left(\frac{Z-\mu}{\sigma} > \frac{\mu + 3\sigma - \mu}{\sigma}\right) \simeq P(Z > 3) \text{ în } N(0,1)$$

pnorm $(3,0,1) = 0.99865$

c) Eșantion de 100:

X_1 avem $\mu = E[X_i]$, $\sigma = \sigma$

$$\overline{X_{100}} = \frac{X_1+X_2+\dots+X_{100}}{100} \Rightarrow E(X_n) = \frac{100\mu}{100} = \mu$$

$$Var(\overline{X_{100}}) = \frac{1}{n^2} n\sigma^2 = \frac{\sigma^2}{n} \Rightarrow \sigma_{X_n} = \frac{\sigma}{\sqrt{n}} = \frac{\sigma}{10}$$

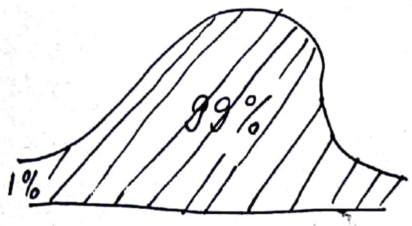
$$P(\overline{X_{100}} > \mu+3\sigma) \overset{cf\,TLC}{\simeq} P\left(\frac{\overline{X_{100}}-\mu}{\frac{\sigma}{100}} > \frac{\mu+3\sigma-\mu}{\frac{\sigma}{10}}\right)$$

$$= P(Z > 30) \text{ în } N(0,1) = 1$$

Deoarece $\overline{X_{100}}$ este aproximativ o distribuție normală cu aceeași medie, dar cu o variație mult mai mică, este normal ca pentru $\overline{X_{100}}$ să avem $P(X > \mu+3\sigma)$ mai mare decât în X_i.

d) Eșantion talie 100 în $N(0,1)$

Ce valoare nu poate fi depășită cu o prob de 99%.



quantila de ordin 0.01

$$P(\overline{X_{100}} < \mu) = 1\%$$

$$\overline{X_{100}} \begin{cases} \mu = 0 \\ \sigma_{\overline{X_{100}}} = \frac{1}{\sqrt{100}} = \frac{1}{10} \end{cases} \simeq N(\mu, \sigma_{\overline{X_{100}}}^2)$$

$$qnorm\left(0.01, 0, \frac{1}{10}\right) = -0.2326$$

e) X_i = o măsurătoare, urmează distribuția $N(0,1)$

50 măsurători

$L_1 = 6$

$L_2 = 13$

$$\overline{X_{50}} = \begin{cases} \mu = 0 \\ \sigma = \frac{1}{\sqrt{50}} \end{cases}$$

Pentru L_2: $\overline{X_{50}} \simeq N\left(0, \frac{1}{\sqrt{50}}^2\right)$

Fie y_1 = valoarea cea mai mică (din L_1)

y_2 = valoarea cea mai mare (din L_2)

$H_2(x) = F(x)^n$ (conform a))

$H_2(13) = \text{pnorm}(13, 10, 1)^{50} = 0.9346$

$\Rightarrow P(Y_2 > 13) = 0.0654$

Există o probabilitate de 93.46 % ca rezultatul să fie greșit și 6.54 % ca rezultatul să fie corect.

Pt valoarea inferioară 6.

$P(Y_1 < 6) = 1 - [1 - F(6)]^{50} = 1 - 0.9984 = 0{,}0016.$

Așadar această valoare e greșită cu o probabilitate mai mare de 99 %.

Exercițiul 5

$$f_{(x,y)}(x,y) = \begin{cases} k(x+y+1), & x \in [0,1], \ y \in [0,2] \\ 0 & , \text{altfel} \end{cases}$$

a) $k = ?$

$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{xy}\,dx\,dy = 1$ a.î. $f_{(x,y)}$ fct de densitate

$$\int_{\infty}^{\infty}\int_{\infty}^{\infty} f_{xy}\,dx\,dy = \int_0^2\int_0^1 (kx + ky + 1\cdot k)\,dx\,dy =$$

$$= \int_0^2 \left(\frac{k}{2}x^2 + kxy + kx\right)\Big|_0^1 dy =$$

$$= \int_0^2 \left(\frac{k}{2} + ky + 1\right)dy = \frac{ky}{2} + \frac{ky^2}{2} + ky \Big|_0^2$$

$= k + 2k + 2k = 5k$

$5k = 1 \Rightarrow k = \frac{1}{5}$

b) Densitățile marginale

$f_X(x) = \int_0^2 f_{XY}(x,y)\,dy = \frac{1}{5}\int_0^2 (x+y+1)\,dy$

$= \frac{1}{5}\left(xy + \frac{y^2}{2} + y\right)\Big|_0^2 = \frac{1}{5}(2x+4) = \frac{2x+4}{5}$

$f_Y(y) = \int_0^1 f_{(X,Y)}(x,y)\,dx = \int_0^1 \frac{1}{5}(x+y+1)\,dx = \frac{1}{5}\left(\frac{x^2}{2} + xy + x\right)\Big|_0^1$

$= \frac{y+3}{10}$

c) X, Y independent

$f_{(X,Y)}(x,y) = f_X(x) \cdot f_Y(y) = \frac{1}{5}(2x+4)\frac{y+3}{10} = \frac{(x+2)(y+3)}{25}$

$= \frac{xy + 3x + 2y + 5}{25} \neq f_{(X,Y)}(x,y) \Rightarrow$ Nu sunt independente

d) $F_X(x) = F_{X,Y}(x,2) = \frac{2x(x+4)}{10}$ (din $\mathcal{F}$)

$F_Y(\mathcal{H}) = F_{XY}(1,\mathcal{H}) = \frac{x(x+3)}{10}$ (din $\mathcal{H}$)

$F_{X,Y}(x,y) = \int_0^y \int_0^x f_{XY}(x,y)\,dx\,dy = \int_0^y \frac{1}{5}\left(\frac{x^2}{2} + xy + x\right)dy =$

$= \frac{1}{5}\left(\frac{yx^2}{2} + \frac{xy^2}{2} + xy\right) = \frac{xy}{2\cdot 5}(x+y+2) = \frac{xy(x+y+2)}{10}$

e) $Z = (X \mid Y = y)$

$\int_0^1 f_{XY}(x,y)\,dx$, cu y fixat

$= \int_0^1 \frac{1}{5}(x+y+1) = \frac{1}{5}\left(\frac{x^2}{2} + yx + x\right)\Big|_0^1 = \frac{3+2y}{10}$

Cum $\frac{3+y\cdot 2}{10}$ este aria de sub. funcția $f(x,y)$ cu y fixat

$$\Rightarrow f_Z(x) = \frac{f(x,y)}{\frac{3+y\cdot 2}{10}} = \frac{\frac{1}{5}(x+y+1)}{\frac{3+y}{10}} = \frac{2(x+y+1)}{3+y\cdot 2}$$

$$f_Z(x) = \begin{cases} \frac{2(x+y+1)}{3+2\cdot y}, & x\in[0,1] \\ 0, & \text{altfel} \end{cases}$$

$R = (Y \mid X = x)$

$$\int_0^2 f_{XY}(x,y)\,dy = \int_0^2 \frac{1}{5}(x+y+1) = \frac{1}{5}\left(xy+\frac{y^2}{2}+y\right)\Big|_0^2$$
$$= \frac{1}{5}(2x+2+2) = \frac{2x+4}{5}$$

$$\Rightarrow f_R(x) = \frac{f(x,y)}{\frac{2x+4}{5}} = \frac{\frac{1}{5}(x+y+1)}{\frac{2x+4}{5}} = \frac{x+y+1}{2x+4}, \ x \text{ fixat}$$

$$\Rightarrow f_R(x) = \begin{cases} \frac{y+x+1}{2x+4}, & x\in[0,2] \\ 0, & \text{altfel} \end{cases}$$

Exercițiul 6

X nr de zile de naștere distincte, $n = 110$ persoane, 365 zile

$$X_i = \begin{cases} 1, & \text{cineva e născut în ziua } i \\ 0, & \text{altfel} \end{cases}$$

$E(X_i) = P$ (cineva e născut în ziua i)
$$= 1 - \left(\frac{364}{365}\right)^{110}$$

$$\Rightarrow E[X] = 365 \times \left[1-\left(\frac{364}{365}\right)^{110}\right] = 365 - \frac{364^{110}}{365^{109}} = 365 - 269 = 96$$

Varianta:

$$Var(X) = Var(X_1+X_2+\ldots+X_{365})$$
$$= Var(X_1)+Var(X_2)+\ldots+Var(365)+\sum_{i,j\in[1,\ldots 365)} Cov(X_i, X_j)$$

Rezultat din formula $Var(X+Y) = Var(X) + Var(Y) + 2Cov(X,Y)$

$$Var(X_1) = Var(X_2) = \dots = Var(365) = E(X_i^2) - E(X_i)^2 =$$

$$= 1^2 \cdot \left[1 - \left(\frac{364}{365}\right)^{110}\right] - \left[1 - \left(\frac{364}{365}\right)^{110}\right]^2$$

$$= 1 - \left(\frac{364}{365}\right)^{110} - 1 + 2\left(\frac{364}{365}\right)^{110} - \left(\frac{364}{365}\right)^{220}$$

$$= \left(\frac{364}{365}\right)^{110} - \left(\frac{364}{365}\right)^{220}$$

$$Cov(X_i, Y_i) = E(X_i, Y_i) - E(X_i)E(X_j)$$

$$E(X_i X_j) = P(X_i = 1, X_j = 1)$$

$$= 1 - P(X_i = 0 \text{ sau } X_j = 0)$$

$$= 1 - P(X_i = 0) - P(X_j = 0) + P(X_i = 0 \text{ și } X_j = 0)$$

$$= 1 - \left(\frac{364}{365}\right)^{110} - \left(\frac{364}{365}\right)^{110} + \left(\frac{363}{365}\right)^{110}$$

$$\Rightarrow Cov(X_i X_j) = 1 - \left(\frac{364}{365}\right)^{110} - \left(\frac{364}{365}\right)^{110} + \left(\frac{363}{365}\right)^{110} -$$

$$- \left(1 - \left(\frac{364}{365}\right)^{110}\right)\left(1 - \left(\frac{364}{365}\right)^{110}\right) = \left(\frac{363}{365}\right)^{110} - \left(\frac{364}{365}\right)^{220}$$

$$\Rightarrow Var(X) = 365 \cdot Var(X_i) + 364 \cdot 365 \cdot \left(\left(\frac{363}{365}\right)^{110} - \left(\frac{364}{365}\right)^{220}\right)$$

$$= 365\left(\frac{364}{365}\right)^{110} - 365\left(\frac{364}{365}\right)^{220} + 364 \cdot 365 \cdot \left(\left(\frac{363}{365}\right)^{110} - \left(\frac{364}{365}\right)^{220}\right)$$

$$= 365\left(\frac{364}{365}\right)^{110} + 364 \cdot 365 \cdot \left(\frac{363}{365}\right)^{110} - 365^2\left(\frac{364}{365}\right)^{220}$$

Exercițiul 7

X_1, X_2 v.a. $N(0,1)$

$Y_1 = \min(X_1, X_2)$

$Y_2 = \max(X_1, X_2)$

$F_{Y_2}(X) = P(\max(X_1, X_2) < X) = P(X_1 < X \text{ și } X_2 < X) = \phi(x)^2$

$\phi(z)$ fct de repartiție a lui $N(0,1)$

$$E[Y_2] = \int_{-\infty}^{\infty} x \, \phi(x)^2 \, dx = \frac{1}{\sqrt{\pi}}$$

$$= \int_{-\infty}^{\infty} x^2 \phi(x)^2 \, dx = 1$$

$$Var[Y_2] = 1 - \left(\frac{1}{\sqrt{\pi}}\right)^2 = \frac{\pi - 1}{\pi} \Rightarrow \sigma_{Y_1} = \sqrt{\frac{\pi-1}{\pi}}$$

$$X_1 + X_2 = Y_1 + Y_2$$

$$\Rightarrow Var(X_1 + X_2) = Var(Y_1 + Y_2) = Var(Y_1) + Var(Y_2) + 2\,Cov(Y_1, Y_2)$$

$$E[X_1 + X_2] = 0$$

$$V(X_1 + X_2) = n\sigma^2 = 2 \cdot 1^2 = 2$$

$$\Rightarrow Cov(Y_1, Y_2) = \frac{2 - \left(1 - \frac{1}{\pi}\right) - \left(1 - \frac{1}{\pi}\right)}{2} = \frac{\frac{2}{\pi}}{2} = \frac{1}{\pi}$$

$$Cor(Y_1, Y_2) = \frac{Cov(Y_1, Y_2)}{\sigma_{Y_1} \sigma_{Y_2}} = \frac{\frac{1}{\pi}}{\sqrt{1 - \frac{1}{\pi}} \sqrt{1 - \frac{1}{\pi}}} = \frac{\frac{1}{\pi}}{\frac{\pi - 1}{\pi}}$$

$$= \frac{1}{\cancel{\pi}} \cdot \frac{\cancel{\pi}}{\pi - 1} = \frac{1}{\pi - 1}$$