Jurusan Teknik Sipil dan Lingkungan
Universitas Gadjah Mada

STATISTIKA

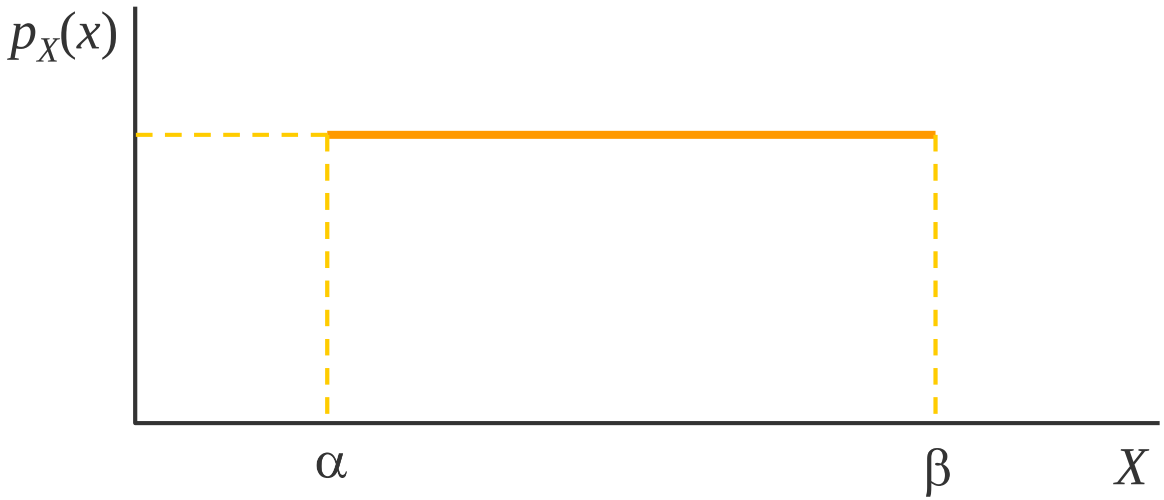
Continuous Probability Distributions

Continuous Probability Distributions

- Normal Distribution
- Uniform Distribution
- Exponential Distribution
- Gamma Distribution
- Lognormal Distribution
- Extreme Value Distributions
  - Extreme Value Type I
  - Extreme Value Type III Minimum (Weibull)
- Beta Distribution
- Pearson Distributions

Normal Distribution

Distribusi Uniform

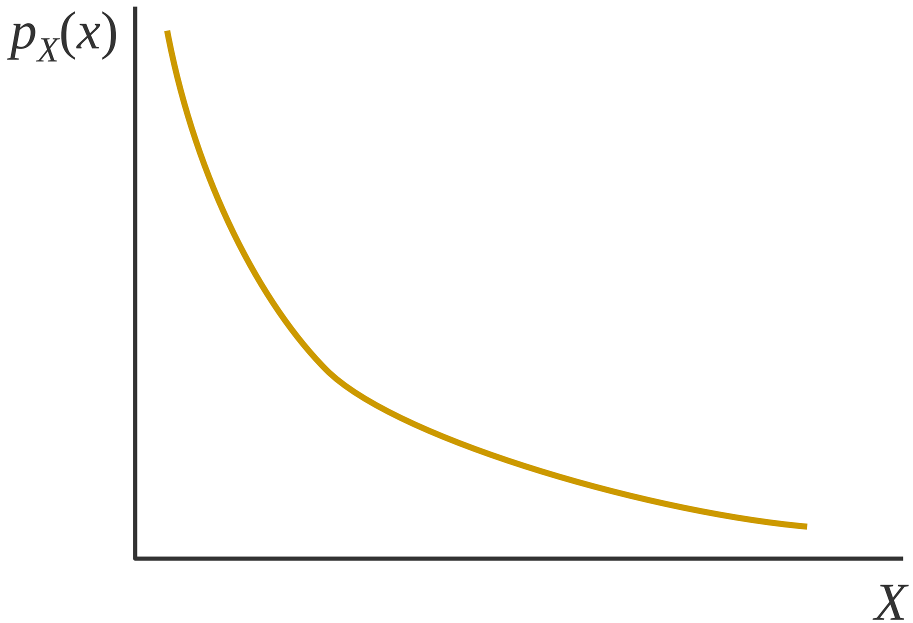


pdf: $p_X\left(x\right) = \dfrac{1}{\beta - \alpha}, \quad \alpha \le x \le \beta$

cdf: $P_X\left(x\right) = \dfrac{x - \alpha}{\beta - \alpha}, \quad \alpha \le x \le \beta$

$E\left(X\right) = \tfrac{1}{2}\left(\beta + \alpha\right) \qquad \mathrm{var}\left(X\right) = \tfrac{1}{12}\left(\beta - \alpha\right)^2$

Distribusi Eksponensial



Coefficient of skew:

$c_s = 2$ (konstan)

Parameter λ:

estimasi $\hat{\lambda} = \dfrac{1}{\bar{X}}$

pdf: $p_X(x) = \lambda\, e^{-\lambda x}, \quad x > 0,\ \lambda > 0$

cdf: $P_X(x) = \int_0^x \lambda\, e^{-\lambda t}\,\mathrm{d}t = 1 - e^{-\lambda x}, \quad x > 0$

Distribusi Gamma

- Penjumlahan sejumlah *n* variabel random berdistribusi exponensial, masing-masing berparameter λ, menghasilkan variabel random berdistribusi gamma dengan parameter λ.

$$\text{pdf} \rightarrow p_X(x) = \lambda^{\eta} x^{\eta-1} e^{-\lambda x} \big/ \Gamma(\eta), \quad x, \lambda, \eta > 0$$

$$\text{cdf} \rightarrow P_X(x) = \int_0^x \lambda^{\eta} t^{\eta-1} e^{-\lambda t} \big/ \Gamma(\eta) \, \mathrm{d}t$$

$$\rightarrow P_X(x) = 1 - e^{-\lambda x} \sum_{j=0}^{\eta-1} \frac{(\lambda x)^j}{j!}, \quad \eta = \text{integer}$$

Distribusi Gamma

$\Gamma(\eta) =$ fungsi gamma

$$\Gamma(\eta) = (\eta - 1)!, \quad \eta = 1, 2, 3, \ldots$$

$$\Gamma(\eta + 1) = \eta\,\Gamma(\eta), \quad \eta > 0$$

$$\Gamma(\eta) = \int_0^\infty t^{\eta - 1} e^{-\eta}\, dt, \quad \eta > 0$$

$$\Gamma(1) = \Gamma(2) = 1$$

$$\Gamma\left(\tfrac{1}{2}\right) = \sqrt{\pi}$$

Mean:

$$E(X) = \eta / \lambda$$

Variance:

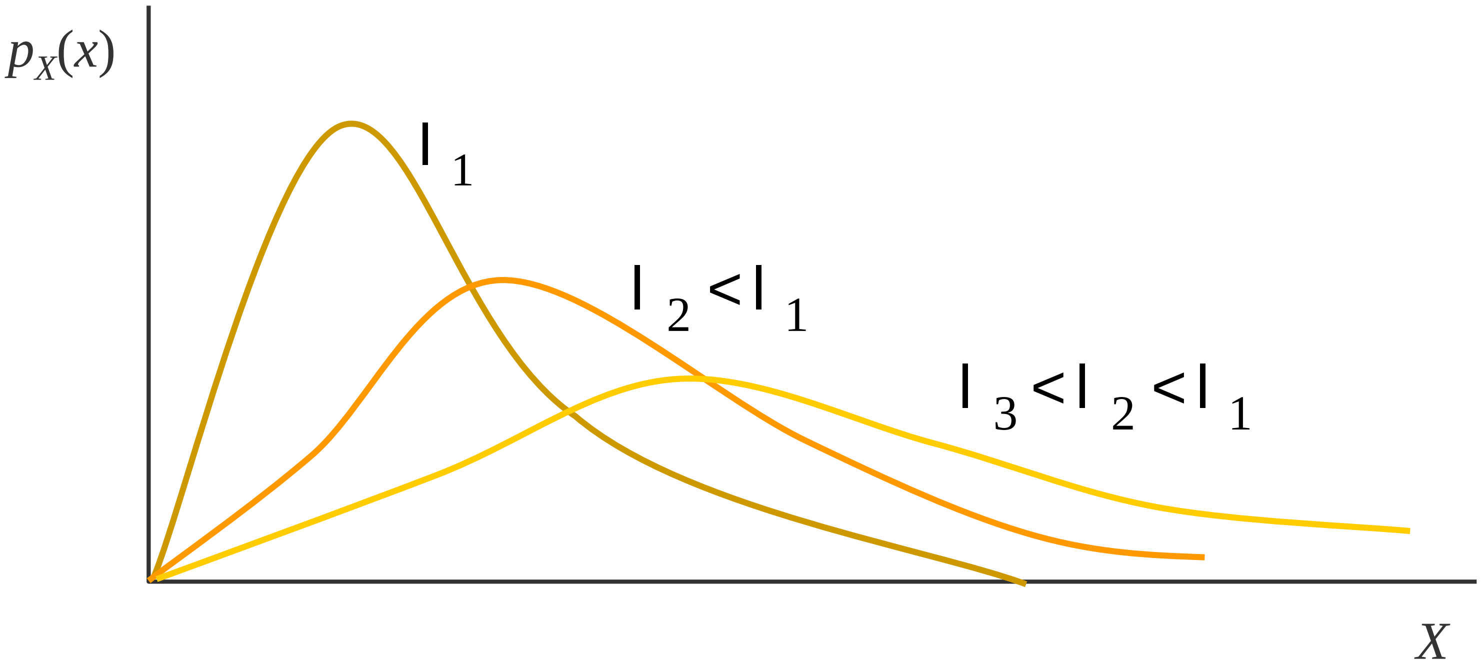
$$\mathrm{var}(X) = \eta / \lambda^2$$

Coef. of skew:

$$\gamma = 2 / \sqrt{\eta}$$

Distribusi Gamma

$\eta = \text{konstan}$



$$p_X(x) = \frac{1}{\Gamma(\eta)} \lambda^{\eta} x^{\eta - 1} e^{-\lambda x}, \quad x, \lambda, \eta > 0$$

Distribusi Lognormal

- Variabel random *X*
  - Jika disusun dari penjumlahan sejumlah pengaruh variabel kecil, maka *X* kemungkinan besar berdistribusi normal.
  - Jika disusun dari perkalian sejumlah pengaruh variabel kecil, maka ln *X* kemungkinan besar berdistribusi normal.

$$X = X_1 + X_2 + \ldots + X_n$$

→ $X_i =$ berdistribusi normal

$X =$ berdistribusi normal

$$X = X_1 \times X_2 \times \ldots \times X_n$$

$$\ln X = \ln X_1 + \ln X_2 + \ldots + \ln X_n$$

→ $\ln X =$ berdistribusi normal

Distribusi Lognormal

$$Y = \ln X$$

$$Y_i = \ln X_i \qquad Y = Y_1 + Y_2 + \ldots + Y_n$$ $\Rightarrow$ Y berdistribusi normal

$$p_Y(y) = \frac{1}{\sigma_Y \sqrt{2\pi}} \, e^{-\frac{1}{2}\left(y - \mu_Y\right)^2 / \sigma_Y{}^2}, \quad -\infty < y < +\infty$$

Distribusi X ?

$$p_X(x) = p_Y(y) \left| \frac{\mathrm{d}\,y}{\mathrm{d}\,x} \right| \qquad Y = \ln X \;\Rightarrow\; \left| \frac{\mathrm{d}\,y}{\mathrm{d}\,x} \right| = \frac{1}{x}, \quad x > 0$$

$$p_X(x) = \frac{1}{x \sigma_Y \sqrt{2\pi}} \, e^{-\frac{1}{2}\left(\ln x - \mu_Y\right)^2 / \sigma_Y{}^2}, \quad x > 0$$

Distribusi Lognormal

Estimasi μ_Y dan σ_Y

$$\left.\begin{array}{l} \mu_Y \rightarrow \bar{Y} \\ \sigma_Y \rightarrow s_Y \end{array}\right\}$$ Data x_i ditransformasikan dulu menjadi $y_i = \ln x_i$

Cara lain:

$$\bar{Y} = \tfrac{1}{2} \ln \left(\frac{\bar{X}^2}{c_v^{\,2} + 1} \right)$$

$$s_Y^{\,2} = \ln\left(c_v^{\,2} + 1\right)$$

c_v = koefisien variasi data asli

$$c_v = \frac{s_X}{\bar{X}}$$

Distribusi Lognormal

- *Mean*

$$E(X) = e^{\left(\mu_Y + \frac{1}{2}\sigma_Y^2\right)}$$

- Varian

$$\operatorname{var}(X) = \mu_X^2 \left(e^{\sigma_Y^2} - 1\right)$$

- Koefisien variasi

$$c_v = \left(e^{\sigma_Y^2} - 1\right)^{\frac{1}{2}}$$

- *Coefficient of skew*

$$\gamma = 3c_v + c_v^3$$

Distribusi Nilai Ekstrem

- Contoh nilai ekstrem
  - Debit banjir
  - Debit minimum
- Nilai-nilai ekstrem variabel random juga merupakan variabel random.
- Distribusi variabel random nilai ekstrem tsb bergantung pada:
  - distribusi variabel random tempat asal variabel nilai ekstrem tsb diperoleh → *parent distribution*
  - jumlah/ukuran sampel

Distribusi Nilai Ekstrem

- Contoh
  - Variabel random

$$X = x_1, x_2, \ldots, x_n$$

Y = nilai ekstrem variabel random tersebut

$$P_Y\left(y\right) = \text{prob}\left(Y \text{ £ } y\right)$$
$$P_{X_i}\left(x\right) = \text{prob}\left(X_i \text{ £ } x\right)$$

$$P_Y\left(y\right) = \text{prob}\left(Y \text{ £ } y\right) = \text{prob}\left(\text{semua } x \text{ yang £ } y\right)$$

Distribusi Nilai Ekstrem

maka:

$$P_Y(y) = P_{X_1}(y) \cdot P_{X_2}(y) \cdots P_{X_n}(y) = \left[P_X(y)\right]^n$$

$$p_Y(y) = \frac{\mathrm{d}\,P_Y(y)}{\mathrm{d}\,y} = n\left[P_X(y)\right]^{n-1} \frac{\mathrm{d}\,P_X(y)}{\mathrm{d}\,y}$$

$$= n\left[P_X(y)\right]^{n-1} p_X(y)$$

Distribusi Nilai Ekstrem

- Contoh
  - Waktu antara 2 hujan berurutan berdistribusi eksponensial.
  - Waktu rata-rata antara 2 hujan = 4 hari
  - Waktu antara tsb merupakan kejadian *independent* satu dengan yang lain
  - Dicari:
    - waktu antara terbesar, misal probabilitas waktu antara tsb lebih besar daripada 8 hari.

Distribusi Nilai Ekstrem

Ditinjau 10 kejadian hujan

$h \quad h \quad h \quad h \quad h \quad h \quad h \quad h \quad h \quad h$

$x_1 \quad x_2 \quad x_3 \quad x_4 \quad x_5 \quad x_6 \quad x_7 \quad x_8 \quad x_9 \rightarrow n = 9$

Distribusi Eksponensial:

$$p_X(x) = \lambda e^{-\lambda x}, \quad x > 0, \lambda > 0$$

$$P_X(x) = 1 - e^{-\lambda x}, \quad x > 0$$

$$E(X) = \lambda^{-1} \qquad \Rightarrow \quad \lambda = \frac{1}{E(X)} \qquad \hat{\lambda} = \frac{1}{\bar{X}}$$

Distribusi Nilai Ekstrem

$$p_X(x) = \frac{1}{4} e^{-\frac{1}{4}x}$$

$$P_X(x) = 1 - e^{-\frac{1}{4}x}$$

$$P_Y(8) = \text{prob}(Y \le 8) = \text{prob}(\text{semua } x \le 8)$$

$$= \left[P_X(8) \right]^9$$

$$= \left(1 - e^{-2}\right)^9$$

$$= 0.271$$

$$\text{prob}(Y > 8) = 1 - 0.271$$

$$= 0.729$$

Distribusi Nilai Ekstrem

- Permasalahan yang sering ditemui adalah bahwa jenis *parent distribution* tidak diketahui.
- Hal ini diatasi dengan
  - ukuran sampel cukup besar, *n* »
  - pemakaian distribusi asimtotis
  - dikenal 3 jenis distribusi asimtotis
    - Type I – *parent distribution unbounded in direction of the desired extreme and all moments of the distribution exist* (*exponential type distributions*)
    - Type II – *parent distribution unbounded in direction of the desired extreme and all moments of the distribution do not exist* (*Cauchy type distributions*)
    - Type III – *parent distribution bounded in the direction of the desired extreme* (*limited distributions*)

Distribusi Nilai Ekstrem

- Permasalahan yang menjadi interest umumnya menyangkut nilai-nilai ekstrem maximum atau extrem minimum.
- Beberapa contoh *parent distributions*
  - Type I – *extreme value largest* – normal, lognormal, eksponential, gamma
  - Type I – *extreme value smallest* – normal
  - Type II – *extreme value largest or smallest* – distribusi Cauchy
  - Type III – *extreme value largest* – distribusi beta
  - Type III – *extreme value smallest* – beta, lognormal, gamma, eksponential

Distribusi Nilai Extrem (9)

- Permasalahan di bidang hidrologi
  - Type II – *extreme value largest or smallest*
    - jarang dijumpai/dipakai
  - Type I – *extreme value largest*
    - nilai ekstrem maksimum sering mengikuti distribusi jenis ini mengingat banyak variabel hidrologi *unbounded* di sisi kanan
  - Type III – *extreme value smallest*
    - nilai ekstrem minimum sering mengikuti distribusi jenis ini mengingat banyak variabel hidrologi *bounded* di sisi kiri oleh nilai nol

Type I Extreme Value Distribution (Gumbel Distribution)

$$p_X(x) = \exp\left\{ \mp (x - \beta)/\alpha - \exp\left[\mp (x - \beta)/\alpha \right] \right\} / \alpha$$

$$-\infty < x < +\infty$$

$$-\infty < \beta < +\infty$$

$$\alpha > 0$$

− untuk nilai maksimum

\+ untuk nilai minimum

α = skala

β = lokasi = mode

Type I Extreme Value Distribution (Gumbel Distribution)

$$E(X) = b + 0.577a \quad \text{(max)}$$

$$= b - 0.577a \quad \text{(min)}$$

$$\operatorname{var}(X) = 1.645a^2 \quad \text{(max/min)}$$

$$g = 1.1396 \quad \text{(max)}$$

$$= -1.1396 \quad \text{(min)}$$

Type I Extreme Value Distribution (Gumbel Distribution)

Dengan memakai transformasi $Y = \dfrac{x - \beta}{\alpha}$

$$p_Y(y) = \exp\left[\mp y - \exp(\mp y)\right]$$

− untuk nilai maksimum

+ untuk nilai minimum

$$P_Y(y) = \int_{-\infty}^{+\infty} \exp\left[\mp t - \exp(\mp t)\right] \mathrm{d}t, \quad -\infty < y < +\infty$$

$$= \exp\left[-\exp(-y)\right] \qquad \text{(max)}$$

$$= 1 - \exp\left[-\exp(y)\right] \qquad \text{(min)}$$

$$P_{\min}(y) = 1 - P_{\max}(-y)$$

Type I Extreme Value Distribution (Gumbel Distribution)

Estimasi parameter α dan β

$$\hat{\alpha} = \frac{s}{1.283}$$

$$\hat{\beta} = \bar{X} - 0.45s \qquad \text{(max)}$$

$$= \bar{X} + 0.45s \qquad \text{(min)}$$

Type II Extreme Value Distribution

$$P_X(x) = 0 \qquad \text{if } x \le \beta$$

$$= \exp\left[-\left(\frac{u-\beta}{x-\beta} \right)^k \right] \qquad \text{if } x \ge \beta$$

$k > 0 \quad \rightarrow \quad$ bentuk

$u - \beta > 0 \quad \rightarrow \quad$ skala

$\beta \quad \rightarrow \quad$ lokasi

$$E(X) = \beta + (u-\beta)\,\Gamma(1 - 1/k)$$

$$\operatorname{var}(X) = (u-\beta)^2 \left[\Gamma(1 - 2/k) - \Gamma^2(1 - 1/k) \right] \qquad k > 2$$

Type III Extreme (Minimum) Value Distribution (Weibull Distribution)

$$p_X(x) = \alpha x^{\alpha-1} \beta^{-\alpha} \exp\left[-\left(\frac{x}{\beta}\right)^{\alpha}\right] \quad x \geq 0$$

$$P_X(x) = 1 - \exp\left[-\left(\frac{x}{\beta}\right)^{\alpha}\right]$$

$$E(X) = \beta\,\Gamma(1+1/\alpha)$$

$$\mathrm{var}(X) = \beta^2 \left[\Gamma(1+2/\alpha) - \Gamma^2(1+1/\alpha)\right]$$

$$\gamma = \frac{\Gamma(1+3/\alpha) - 3\Gamma(1+2/\alpha)\Gamma(1+1/\alpha) + 2\Gamma^3(1+1/\alpha)}{\left[\Gamma(1+2/\alpha) - \Gamma^2(1+1/\alpha)\right]^{3/2}}$$

Type III Extreme (Minimum) Value Distribution (Weibull Distribution)

Estimates:

$$\lambda = \beta^{-\alpha}$$

$$\hat{\lambda} = \frac{n}{\sum_{i=1}^{n} x_i^{\hat{\alpha}}}$$

$$\hat{\alpha} = \frac{n}{\hat{\lambda} \sum_{i=1}^{n} x_i^{\hat{\alpha}} \ln x_i - \sum_{i=1}^{n} \ln x_i}$$

$$\hat{\beta} = \left(\hat{\alpha}\right)^{-1/\hat{\alpha}}$$

Extreme Value Distributions

- Silakan baca discussion pada hlm 118 (Haan, 1982)

Beta Distribution

Distribusi yang memiliki batas atas dan batas bawah

$$p_X(x) = x^{\alpha-1}(1-x)^{\beta-1} \big/ \mathrm{B}(\alpha,\beta), \quad 0 < x < 1 \text{ and } \alpha,\beta > 0$$

$$\mathrm{B}(\alpha,\beta) = \text{beta function}$$

$$= \int_0^1 x^{\alpha-1}(1-x)^{\beta-1}\,dx$$

$$= \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$$

$$E(X) = \frac{\alpha}{\alpha+\beta}$$

$$\operatorname{var}(X) = \frac{\alpha\beta}{(\alpha+\beta+1)(\alpha+\beta)^2}$$

Pearson Type III Distribution

$$p_X(x) = p_0 \left(1 + x/\alpha\right)^{\alpha/\delta} e^{-x/\delta}$$

- mode di $x = 0$
- batas bawah di $x = -\alpha$

Dengan transformasi (translasi) sehingga:

- mode di $x = \alpha$
- batas bawah di $x = 0$

$$p_X(x) = p_0\, e^{-(x-\alpha)/\delta} \left(\frac{x}{\alpha}\right)^{\alpha/\beta}$$

Distribusi Chi-square

Distribusi t

Distribusi F

DISTRIBUSI SAMPEL STATISTIK

Chi-square Distribution

$Z = \dfrac{X - \mu}{\sigma}$ variabel random berdistribusi normal

$Y = \sum_{i=1}^{n} Z_i^{2}$ berdistribusi chi-square dengan *n* degrees of freedom

Distribusi chi-square = distribusi gamma dengan
λ = ½
η = kelipatan ½

$$p_{\chi^2}(x) = \frac{x^{-(1-\nu/2)}\, e^{-x/2}}{2^{\nu/2}\,\Gamma(\nu/2)} \qquad x, \nu > 0$$

$$\nu = 2\eta$$

$$E\left(\chi^2\right) = \nu \qquad \mathrm{var}\left(\chi^2\right) = 2\nu \qquad \hat{\nu} = \bar{X}$$

t Distribution

Y = normal standar

U = chi-square

} Y dan U independent

$$X = Y\frac{\sqrt{\nu}}{\sqrt{U}}$$

→ berdistribusi t dengan ν degrees of freedom

$$p_T(t) = \frac{\Gamma\left[\frac{1}{2}(\nu+1)\right]\left(1+t^2/\nu\right)^{-\frac{1}{2}(\nu+1)}}{\sqrt{\pi\nu}\,\Gamma(\nu/2)}$$

$$-\infty > t < +\infty$$

$$\nu > 0$$

$$E(T) = 0$$

$$\mathrm{var}\left(\chi^2\right) = \frac{\nu}{\nu-2} \quad \text{untuk } \nu > 2$$

F Distribution

U = chi-square dengan $\gamma = m$ *degrees of freedom*
V = chi-square dengan $\gamma = n$ *degrees of freedom*
} U dan V *independent*

maka:

$X = (U/m)(V/n)$ → berdistribusi F dengan $\gamma_1 = m$ dan $\gamma_2 = n$ *degrees of freedom*

$$p_F(f) = \frac{\Gamma\left[\frac{1}{2}(\gamma_1+\gamma_2)\right]\gamma_1^{\gamma_1/2}\,\gamma_2^{\gamma_2/2}\,f^{\frac{1}{2}(\gamma_1-2)}}{(\gamma_2+\gamma_1 f)^{\frac{1}{2}(\gamma_1+\gamma_2)}\,\Gamma(\gamma_1/2)\,\Gamma(\gamma_2/2)} \qquad \gamma_1, \gamma_2 > 0$$

$$E(F) = \frac{\gamma_1}{\gamma_2 - 2} \qquad \mathrm{var}(F) = \frac{\gamma_2^{\,2}(\gamma_1+2)}{\gamma_1(\gamma_2-2)(\gamma_2-4)}$$

Terima kasih