パターン認識と機械学習

2.3.7 スチューデントのt分布

吉田拓夢

周辺分布

$$
\begin{aligned}
p(x|\mu,a,b) &= \int_0^{\infty} \mathcal{N}(x|\mu,\tau^{-1})\mathrm{Gam}(\tau|a,b)d\tau \\
&= \int_0^{\infty} \frac{b^a e^{(-b\tau)}\tau^{a-1}}{\Gamma(a)} \left(frac\tau 2\pi \right)^{1/2} \exp\left\{ -\frac{\tau}{2}(x-\mu)^2 \right\} d\tau \\
&= \frac{b^a}{\Gamma(a)} (\frac{1}{2\pi})^{1/2} \left[b + \frac{(x-\mu)^2}{2} \right]^{-a-1/2} \Gamma(a+1/2)
\end{aligned}
$$

- 1変数ガウス分布 $\mathcal{N}(x|\mu,\tau^{-1})$

→精度の事前分布をガンマ分布 $\mathrm{Gam}(\tau|a,b)$

精度を積分消去し、変数置換 $z = \tau[b + (x-\mu)^2/2]$

スチューデントのt分布

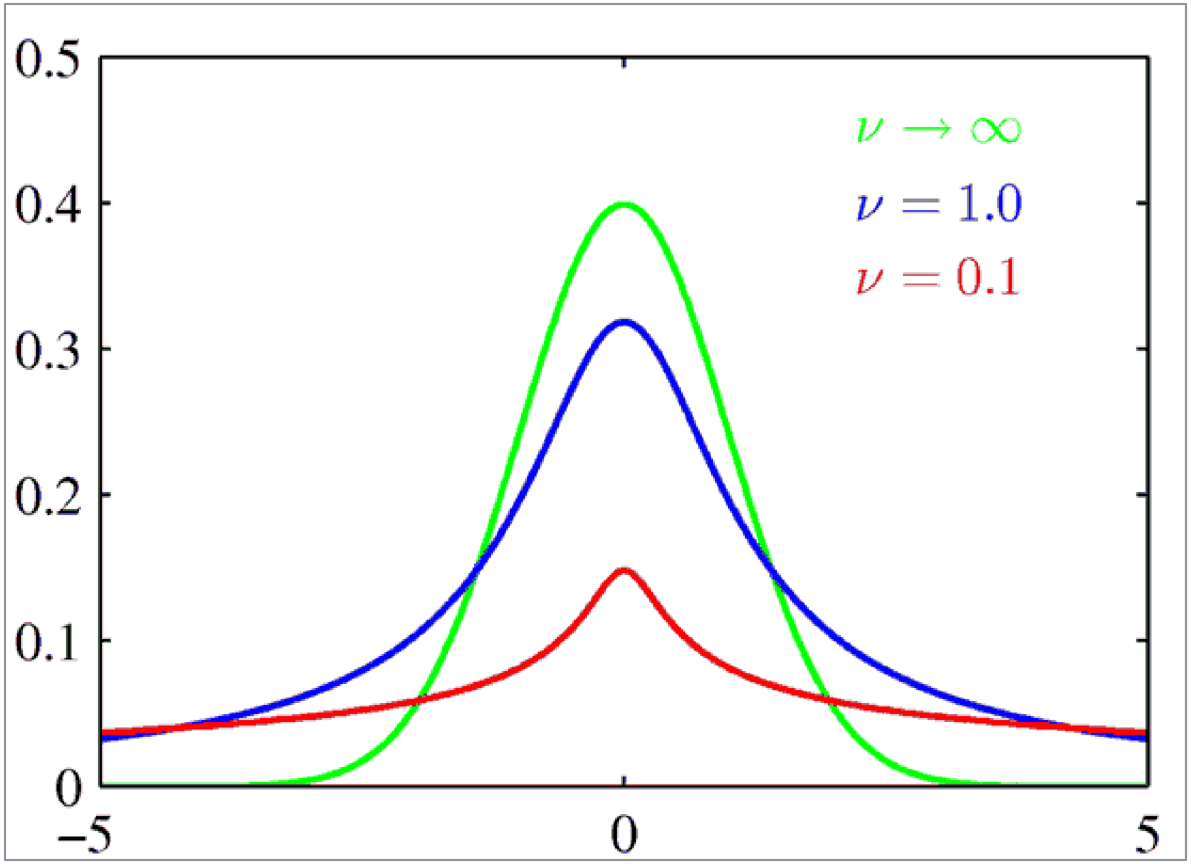
- パラメータを定義：ν=2a,λ=a/b
  ν：自由度
  λ：精度(必ずしも分散の逆数ではない)

$$\mathrm{St}(x|\mu,\lambda,\nu)=\frac{\Gamma(\nu/2+1/2)}{\Gamma(\nu/2)}\left(\frac{\lambda}{\pi\nu}\right)^{1/2}\left[1+\frac{\lambda(x-\mu)^2}{\nu}\right]^{-\nu/2-1/2}$$

- 更にη=τb/aと置き換え

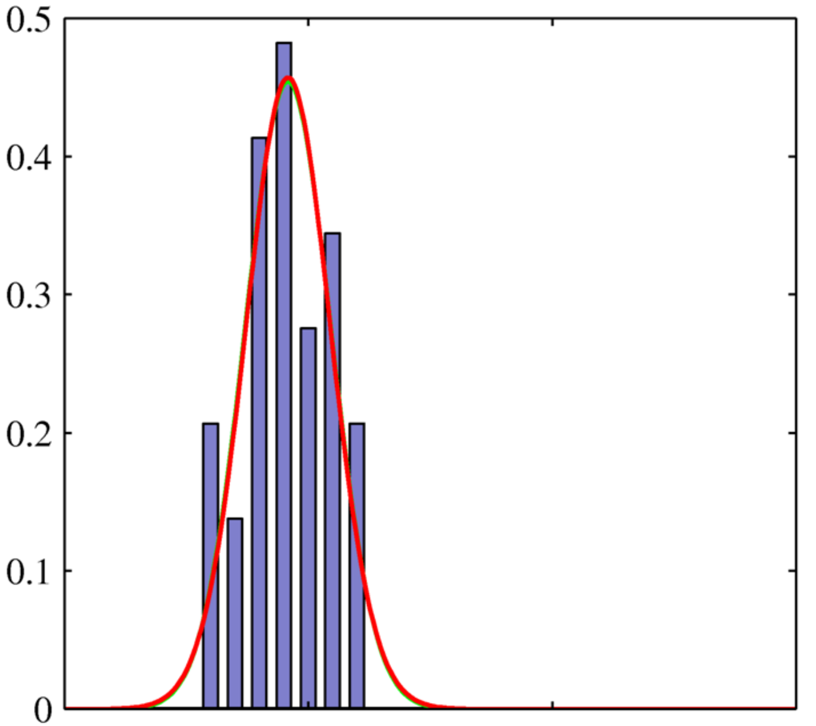
$$\mathrm{St}(x|\mu,\lambda,\nu)=\int_0^{\infty}\mathcal{N}(x|\mu,(\eta\lambda)^{-1})\mathrm{Gam}(\eta|\nu/2,\nu/2)d\eta$$

- スチューデントのt分布

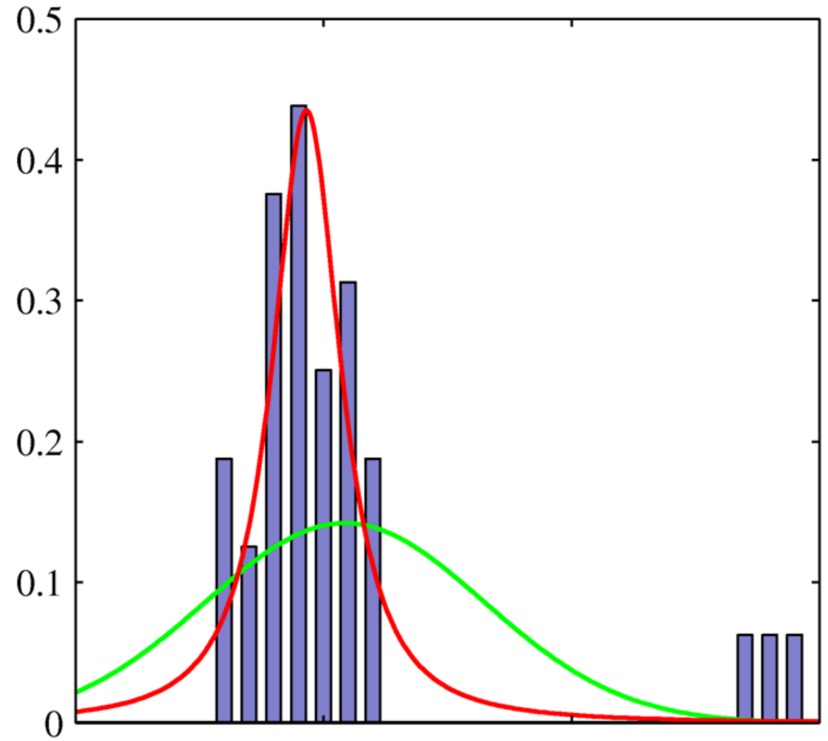


- ν→∞の極限でガウス分布
- ν=1でコーシー分布

- 頑健性の比較(赤:ガウス分布　緑:t分布)



(a)



(b)

多変量の場合

$$\mathrm{St}(\mathbf{x}|\boldsymbol{\mu}, \boldsymbol{\Lambda}, \nu) = \int_0^\infty \mathcal{N}(\mathbf{x}|\boldsymbol{\mu}, (\eta\boldsymbol{\Lambda})^{-1})\mathrm{Gam}(\eta|\nu/2, \nu/2)d\eta$$

$$\mathrm{St}(\mathbf{x}|\boldsymbol{\mu}, \boldsymbol{\Lambda}, \nu) = \frac{\Gamma(D/2+\nu/2)}{\Gamma(\nu/2)}\frac{|\boldsymbol{\Lambda}|^{1/2}}{(\pi\nu)^{D/2}}\left[1+\frac{\Delta^2}{\nu}\right]^{-D/2-\nu/2}$$

→Dはxの次元数、Δはマハラノビス距離

$$\Delta^2 = (\mathbf{x}-\boldsymbol{\mu})^{\mathrm{T}}\boldsymbol{\Lambda}(\mathbf{x}-\boldsymbol{\mu})$$

多変量の場合

- 1変数の結果に対応した、以下の性質を持つ

$$
\begin{aligned}
\mathbb{E}[\mathbf{x}] &= \boldsymbol{\mu}, & \nu > 1 \text{ のとき} \\
\mathrm{cov}[\mathbf{x}] &= \frac{\nu}{(\nu - 2)}\mathbf{\Lambda}^{-1}, & \nu > 2 \text{ のとき} \\
\mathrm{mode}[\mathbf{x}] &= \boldsymbol{\mu}
\end{aligned}
$$