

$$\pi(\theta/D) = \frac{f(D/\theta)}{P(D)} \pi(\theta) \qquad P(D) = \int f(D/\theta) \pi(\theta) d\theta$$

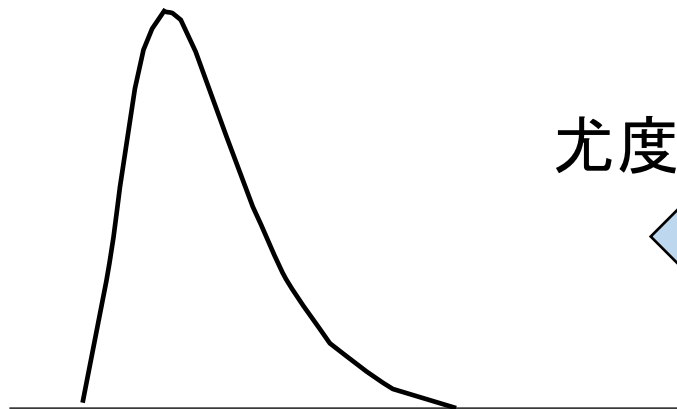
平均値 $M = \frac{\int \theta f(D/\theta) \pi(\theta) d\theta}{\int f(D/\theta) \pi(\theta) d\theta}$

分散 $S^2 = \frac{\int (\theta - M)^2 f(D/\theta) \pi(\theta) d\theta}{\int f(D/\theta) \pi(\theta) d\theta}$

$$\pi(\theta/D) = k f(D/\theta) \pi(\theta)$$

事後分布 = 定数 $\times$ 尤度 $\times$ 事前分布

事前分布と同種類の分布

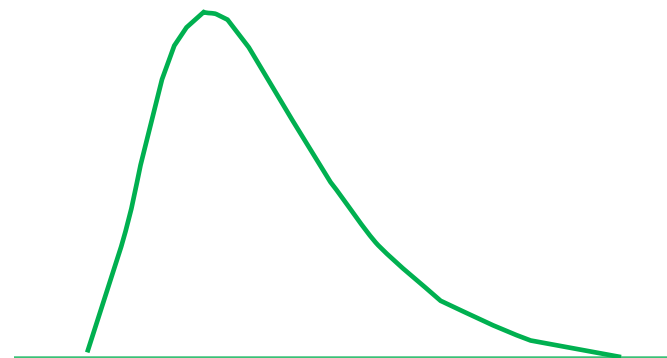


事後分布

尤度を掛ける



自然共役分布



事前分布

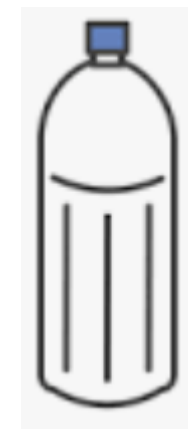


積分計算を省ける

平均値 $M = \frac{\int \theta f(D/\theta) \pi(\theta) d\theta}{\int f(D/\theta) \pi(\theta) d\theta}$

分散 $S^2 = \frac{\int (\theta - M)^2 f(D/\theta) \pi(\theta) d\theta}{\int f(D/\theta) \pi(\theta) d\theta}$

製造される内容量200mLの清涼飲料水のペットボトルの内容量 x は正規分布に従い、分散は1である。3本サンプリングしたところ、201、202及び203mLであった。この工場で製造のペットボトルの内容量 x の平均値 θ の確率分布を求めよ



正規分布なので、 $f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{(x-\theta)^2}{2}}$

尤度 $f(D/\theta)$ は、データの値を掛け合わせた数値なので

$$\begin{aligned} f(D/\theta) &= \frac{1}{\sqrt{2\pi}} e^{-\frac{(201-\theta)^2}{2}} \times \frac{1}{\sqrt{2\pi}} e^{-\frac{(202-\theta)^2}{2}} \times \frac{1}{\sqrt{2\pi}} e^{-\frac{(203-\theta)^2}{2}} \\ &= \frac{1}{\sqrt{2\pi}} e^{-\frac{(\theta-202)^2}{2 \times \frac{1}{3}}} \end{aligned}$$

事前分布 $\pi(\theta)$ は不明なので、平均値200mL、分散4としておきます。

$$\pi(\theta) = \frac{1}{2\sqrt{2\pi}} e^{-\frac{(x-200)^2}{2 \times 4}}$$

← 自然共役分布

$$\pi(\theta/D) = k f(D/\theta) \pi(\theta) \propto e^{-\frac{(\theta-202)^2}{2 \times \frac{1}{3}}} \times \frac{1}{2\sqrt{2\pi}} e^{-\frac{(x-200)^2}{2 \times 4}} \propto e^{-\frac{\left(x - \frac{2624}{13}\right)^2}{2 \times \frac{4}{13}}}$$

事後分布

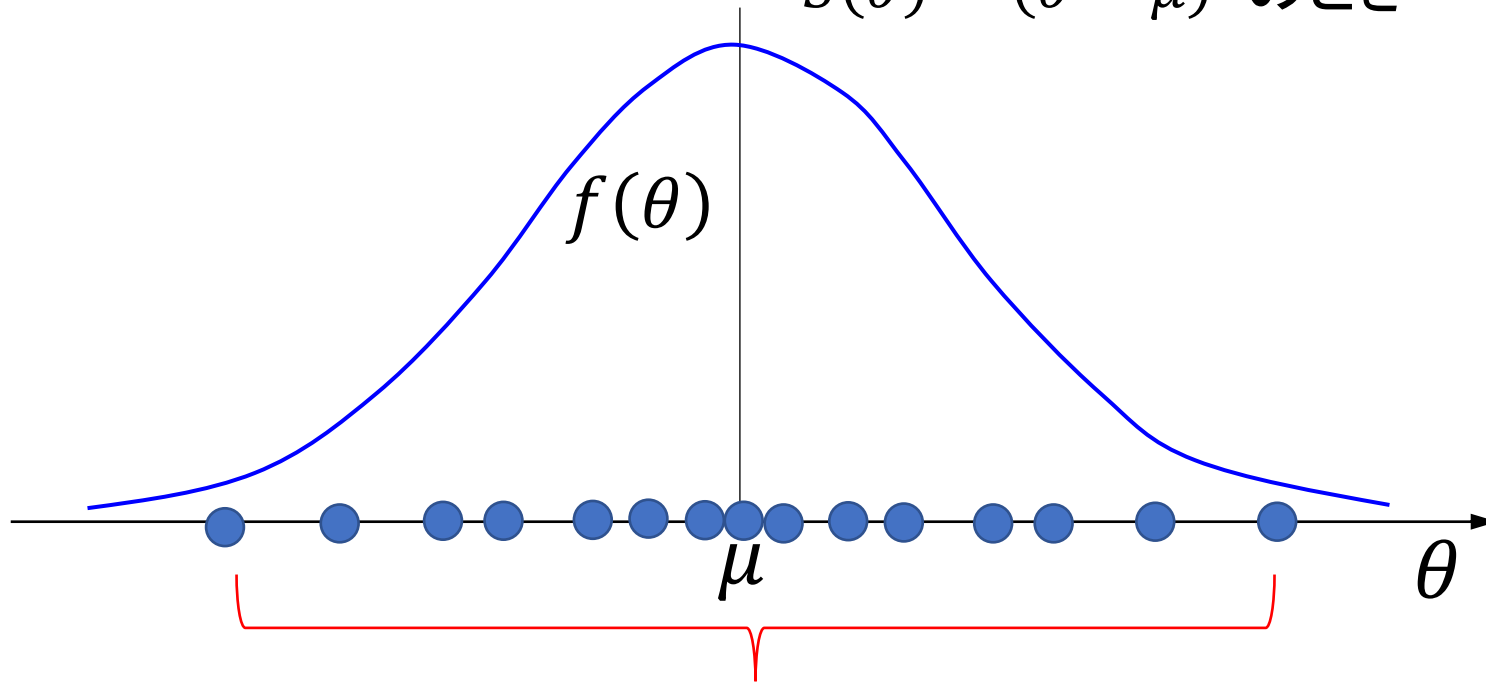
平均値 $\frac{2624}{13} = 202.846 \dots$ 、分散 $\frac{4}{13}$ の正規分布に従う

MCMC (Markov chain Monte Carlo methods) 法

$$S(\theta) \text{の期待値} = \int S(\theta) f(\theta) d\theta$$

$S(\theta) = \theta$ のとき $\rightarrow$ 確率分布の平均値 μ

$S(\theta) = (\theta - \mu)^2$ のとき $\rightarrow$ 確率分布の分散



$f(\theta)$ に従う n 個の乱数 $\theta_1, \theta_2, \dots, \theta_n$

$$\int S(\theta) f(\theta) d\theta \doteq \frac{1}{n} \{S(\theta_1) + S(\theta_2) + \dots + S(\theta_n)\}$$

Red arrows point from the terms $S(\theta_1)$, $S(\theta_2)$, and $S(\theta_n)$ in the equation to the corresponding sample points θ_1 , θ_2 , and θ_n on the horizontal axis of the graph above.

平均値150、分散 10^2 の正規分布に従う乱数1,000個を用いて、
 $S(\theta) = \theta = \text{平均値}\mu$ 、 $S(\theta) = (\theta - \mu)^2 = \text{分散}$ となるかを確認せよ

140.3598	145.4676	151.8084	155.3967	158.3797	150.3211	153.2008	150.1932	161.0188	154.5961
135.5782	157.7267	154.0943	146.8889	152.1034	170.8997	151.6221	141.4394	157.6471	158.2362
151.6249	144.5815	144.5809	150.2186	142.84	137.7217	160.8298	141.5247	143.4291	147.9572
126.1562	167.7001	148.3384	170.3641	139.3817	140.3048	138.494	156.2878	127.4527	122.8237
156.6072	140.9876	145.9023	144.2895	156.2203	150.008	170.2826	153.9697	161.5749	153.0085
153.0259	137.2059	150.3777	156.6665	161.2914	169.2684	154.2295	149.8626	138.1364	128.362
134.212	149.5729	154.4816	160.8497	141.1354	148.9561	154.1043	144.5313	160.3776	155.5816
}									
144.8474	138.8213	134.0827	150.6062	157.364	152.0801	125.9826	146.0186	155.4573	158.8746
152.3017	149.1993	138.1412	151.5306	173.9126	137.0803	154.0151	168.2569	138.1897	154.03
140.1047	152.9589	166.8598	166.1217	165.7093	134.1109	155.3366	159.3185	165.9836	144.7137
144.8616	156.1714	142.6106	130.3164	138.3129	146.1862	134.6939	142.2342	163.6979	130.2686
153.3569	158.0745	147.9113	146.2279	138.668	144.9218	151.012	156.268	151.1359	144.881

$$\begin{aligned}
 S(\theta) &= \theta、\quad \frac{1}{1000}\{S(\theta_1) + S(\theta_2) + \cdots \cdots + S(\theta_{1000})\} = \textcolor{blue}{149.3} = \mu \\
 S(\theta) &= (\theta - \mu)^2、\quad \frac{1}{1000}\{(\theta_1 - \mu)^2 + (\theta_2 - \mu)^2 + \cdots \cdots + (\theta_{1000} - \mu)^2\} = \textcolor{blue}{104.7} = \text{分散}
 \end{aligned}$$