**Bayesowski Model Normalnej Regresji Liniowej – ćwiczenie**

1. Zapoznaj się z danymi zestawionymi w pliku *Wyplaty\_gotowki.xlsx*
2. Niech narzędziem modelowania zależności pomiędzy wartością wypłat gotówki w bankomatach w Polsce a liczbą bankomatów raz wydanych kart płatniczych będzie model regresji postaci:

$W\_{t}=β\_{0}+β\_{1}t+\sum\_{i=2}^{4}β\_{i}Q\_{ti}+β\_{5}B\_{t}+β\_{6}K\_{t}+ε\_{t}$, $ε\_{t}\~iiN(0, τ^{-1})$,

gdzie $Q\_{ti} (i=2, 3, 4)$ reprezentują zero-jedynkowe zmienne sezonowe odpowiadające kwartałom II $(i=2)$, III $(i=3)$ i IV $(i=4)$. Zbuduj stosowne macierze $y$ oraz $X$, odpowiadające macierzowemu zapisowi tego modelu: $y=Xβ+ε$

Dla powyższego modelu przeprowadź estymację bayesowską przyjmując gamma-normalny rozkład *a priori* z zależnością pomiędzy $β$ i $τ$, przy założeniu, że:

* + $E\left(τ\right)=1, Var\left(τ\right)=0,5$
	+ $E\left(β\right)=0\_{(k×1)}$ i $V\left(β\right)=100I\_{k}$ (zatem $Var\left(β\_{i}\right)=100$, $i=1, 2,…, k$, oraz $Cov\left(β\_{i},β\_{j}\right)=0$, $i\ne j$)

Dla poszczególnych parametrów (tak strukturalnych, jak i precyzji składników losowych) wyznacz następujące **charakterystyki** (te w brzegowych rozkładach *a posteriori* zestaw z tymi *a priori*):

 - wartości oczekiwane, mediany, modalne

 - odchylenie standardowe

 - przedziały najwyższej gęstości *a posteriori* (dla poziomu prawdopodobieństwa 0,95)

Dodatkowo, dla parametru odzwierciedlającego związek pomiędzy liczbą bankomatów w wartością wypłat gotówkowych (i tylko dla niego) **skonstruuj wykres** przedstawiający gęstości rozkładów *a posteriori* i *a priori* (na tym samym wykresie).

**Uwaga 2:** W celu ustalenia zakresu wartości na osi OX tego wykresu, możesz przyjąć regułę, że zakres ten stanowi przedział $\left(E\left(y\right)-4D\left(y\right), E\left(y\right)+4D\left(y\right)\right)$, z końcami zaokrąglonymi do np. 2-3 miejsc po przecinku. Przedział ten można następnie podzielić np. na 100 lub więcej równych odcinków, w ten sposób wyznaczając w arkuszu „siatkę punktów” (w postaci kolumny) wartości parametru, w których zostaną obliczone wartości funkcji gęstości rozkładów *a posteriori* i *a priori* (w centrum wykresu będzie się znajdował rozkład *a posteriori*, zaś to, na ile rozkład *a priori* będzie widoczny na tym wykresie, zależeć będzie od tego, jak daleko jest on ulokowany w stosunku do rozkładu *a posteriori*). W każdym punkcie tej „siatki” należy obliczyć wartości funkcji gęstości stosownych rozkładów.

**Uwaga 3**: O ile wartości funkcji gęstości rozkładu gamma można obliczyć w MS Excel za pomocą dostępnej tam funkcji, o tyle nie ma w arkuszu funkcji obliczającej wartość gęstości 1-wymiarowego rozkładu t-Studenta $St(m, p,ν)$, gdzie $m$ oznacza parametr niecentralności (modalną), $p$ – precyzję, zaś $ν$ – liczbę stopni swobody ($ν>0$). Dlatego wartości funkcji gęstości tego rozkładu należy obliczyć wprowadzając „ręcznie” do komórki wzór:

$f\_{St}\left(m, p,ν\right)=\frac{Γ\left(\frac{ν+1}{2}\right)\sqrt{p}}{Γ\left(\frac{ν}{2}\right)\sqrt{νπ}}\left(1+\frac{p\left(x-m\right)^{2}}{ν}\right)^{-\frac{ν+1}{2}}$, $x\in R$

1. Zweryfikuj statystyczną istotność poszczególnych parametrów strukturalnych.