Gibbsian conditioning

Let $(\Omega, \mathcal{F}, \mathbb{P})$ a probability space on which we consider a sequence of iid discrete random variables $(X_n)_{n\geqslant 1}$ with law $\mathbb{P}(X_n = \pm 1) = 1/2$. Let $M_n = (X_1 + \cdots + X_n)/n$ the empirical mean of the first n variables. Fix $\varepsilon > 0$ and $m \in]-1, 1[$ and let B_n be the set $B_n = \{\omega \in \Omega : M_n(\omega) \in [m, m+\varepsilon]\}$. We want to study the limit law as $n \to \infty$ of the k-ple $(X_1, \ldots, X_k)$ when $(X_1, \ldots, X_n)$ is conditioned on the event B_n. More precisely fix $k \geqslant 1$ and let μ_n be the law of $(X_1, \ldots, X_k)$ conditional on B_n:

$$\mu_n(x_1, \ldots, x_k) = \mathbb{P}(X_1 = x_1, \ldots, X_k = x_k | B_n) = \frac{\mathbb{P}(X_1 = x_1, \ldots, X_k = x_k, B_n)}{\mathbb{P}(B_n)}.$$

Our aim is to prove that the family $\{\mu_n\}_{n\geqslant 1}$ converge weakly to the law ρ_λ^k on $\{-1, 1\}^k$ for which all the components are independent and

$$\rho_\lambda^k(x_1, \ldots, x_k) = \rho_\lambda(x_1) \cdots \rho_\lambda(x_k)$$

where ρ is the discrete probability on $\{-1, 1\}$ given by

$$\rho_\lambda(x) = \frac{e^{\lambda x}}{e^\lambda + e^{-\lambda}} \qquad \text{for } x = -1, 1$$

where $\lambda \in \mathbb{R}$ is a fixed number which is determined by the fact that the mean of the measure ρ_λ should be m:

$$m = \sum_{x=-1,1} x \rho_\lambda(x) = \frac{e^\lambda - e^{-\lambda}}{e^\lambda + e^{-\lambda}} = \tanh(\lambda).$$

To prove this weak convergence result you need to understand the discussion on “Gibbsian conditioning” in the “Poly 4” of the lecture notes and proceed as follows:

a) Start by proving the statement for $k = 1$. Note that for any continuous function

$$\sum_{x_1=\pm 1} f(x_1)\, \mu_n(x_1) = \frac{\mathbb{E}[f(X_1) 1_{B_n}]}{\mathbb{E}[1_{B_n}]} = \frac{\mathbb{E}[L_n(f) 1_{B_n}]}{\mathbb{E}[1_{B_n}]}$$

where $L_n(f) = \frac{1}{n} \sum_{i=1}^n f(X_i)$ is the mean of f with respect to the empirical measure of the random vector $(X_i)_{1\leqslant i\leqslant n}$. Observe also that $B_n = \{L_n(h) \in [m, m+\varepsilon]\}$ where $h: \{-1, +1\} \to \mathbb{R}$ is the identity function given by $h(x) = x$. Then

$$\sum_{x_1=\pm 1} f(x_1)\, \mu_n(x_1) = \frac{\mathbb{E}[L_n(f) 1_{L_n(h)\in[m,m+\varepsilon]}]}{\mathbb{E}[1_{L_n(h)\in[m,m+\varepsilon]}]}$$

Use Sanov theorem (Theorem 7 of Poly 4) and Proposition 1 and Corollary 2 of Poly 4 to deduce a large deviation principle for μ_n. Conclude that $\mu_n \to \rho_\lambda$.

b) Follow the discussion on “Gibbsian conditioning” in the “Poly 4” of the lecture notes to extend the argument to $k > 1$. For the purpose of the exam it is enough to prove the statement for $k = 2$: that is we want to prove that conditionally on B_n, the pair (X_1, X_2) coverge weakly to a pair of independent variables each of them with law ρ_λ. (with λ depending on m).