Task 1

$$P\left[X \geq \epsilon\right] = 1 - F(\epsilon) \leq e^{-2m\epsilon^2} \Rightarrow \int_\epsilon^\infty p(t)dt \leq e^{-2m\epsilon^2}$$

Учитывая, что мы хотим оценить $\mathbb{E}\left[e^{2(m-1)X^2}\right] = \int_0^\infty e^{2(m-1)t^2} p(t)dt$, которое достигнет максимума в случае, когда $\int_\epsilon^\infty p(t)dt = e^{-2m\epsilon^2}$. Тогда

$$p(\epsilon) = 4m\epsilon e^{-2m\epsilon^2}$$
$$\mathbb{E}\left[e^{2(m-1)X^2}\right] \leq \int_0^\infty e^{2(m-1)t^2} p(t)dt = \int_0^\infty 4mte^{-2mt^2} e^{2(m-1)t^2} dt =$$
$$= \int_0^\infty 4mte^{-t^2} dt = 2m$$

Task 2

1) Так как P - равномерное распределение, а $Q = \begin{cases} 1 & \text{для } h = h_s, \\ 0 & \text{иначе.} \end{cases}$

Учитывая

$$L_D(Q) \leq L_S(Q) + \sqrt{\frac{D(Q\|P) + \ln m/\delta}{2(m-1)}}$$

получаем

$$L_D(h_s) \leq L_S(h_s) + \sqrt{\frac{\ln \frac{1}{1/|H|} + \ln m/\delta}{2(m-1)}}.$$

2) Воспользуемся неравенством Хефдинга:

$$P[L_D(h) - L_S(h) \geq \varepsilon] \leq 2\exp\left(\frac{-m\varepsilon^2}{2M^4}\right) \Rightarrow$$
$$\forall h' \in H \quad P[L_D(h') - L_S(h') \geq \varepsilon] \leq \sum_{h \in H} P[L_D(h) - L_S(h) \geq \varepsilon] \leq$$
$$\leq 2|H|\exp\left(\frac{-m\varepsilon^2}{2M^4}\right) = \delta \Rightarrow$$
$$\varepsilon = M\sqrt{\frac{\ln|H| + \ln 2/\delta}{2m}} \Rightarrow$$
$$P\left[L_D(h) \leq L_S(h) + M\sqrt{\frac{\ln|H| + \ln 2/\delta}{2m}}\right] \geq 1 - \delta.$$

Получаем, что бОльший размер выборки гарантирует лучшее обобщение и для одинаковоой эмпирической ошибки, чем меньше размер H, тем лучше.

Task 3

1) $|\Delta(h_1) - \Delta(h_2)| \leq |L_D(h_1) - L_D(h_2)| + |L_S(h_1) - L_S(h_2)| = \left|\mathbb{E}\left[(h_1 - y)^2 - (h_2 - y)^2\right]\right| +$

$+ \left|\frac{1}{m}\sum_{i=1}^m \left((h_1(x_i) - y_i)^2 - (h_2(x_i) - y_i)^2\right)\right|$

$(h_1 - y)^2 - (h_2 - y)^2 = (h_1 - h_2)(h_1 + h_2 - 2y) =$
$= (h_1 - h_2)((h_1 - y) + (h_2) - y)) \le 2M \left\| h_1 - h_2 \right\|_\infty$
Тогда
$|\Delta(h_1) - \Delta(h_2)| \le 4M \left\| h_1 - h_2 \right\|_\infty$

2) $P[\underset{h\in H}{sup}|\Delta(h)| \ge \varepsilon] = P[\bigvee_{i=1}^{k} \underset{h\in B_i}{sup}|\Delta(h)| \ge \varepsilon] \le \sum\limits_{i=1}^{k} P[\underset{h\in B_i}{sup}|\Delta(h)| \ge \varepsilon]$

3) Учитывая 1) и то, что радиусы шаров равны $\frac{\varepsilon}{8M}$:
$|\Delta(h_1) - \Delta(h_2)| \le 4M \left\| h_1 - h_2 \right\|_\infty \le \frac{\varepsilon}{2}$
Если для некоторого $h \in B_i$ выполняется $|\Delta(h)| \ge \varepsilon$, тогда $|\Delta(h_i)| \ge \frac{\varepsilon}{2}$, т.е. $P[\underset{h\in B_i}{sup}|\Delta(h)| \ge \varepsilon] \le P[|\Delta(h_i)| \ge \frac{\varepsilon}{2}]$

4) Так как $\Delta(h) = L_D(h) - L_S(h)$, $L_D(h) = \mathbb{E}[(h(x) - y)^2]$,
$L_S(h) = \frac{1}{m}\sum\limits_{i=1}^{m}(h(x_i) - y_i)^2$. Тогда неравенство Хефдинга можно применить к последовательности случайных величин $Z_i = (h(x_i) - y_i)^2$:

$$P[|\Delta(h)| \ge \tfrac{\varepsilon}{2}] \le 2\exp\left(\tfrac{-m\varepsilon^2}{2M^4}\right).$$

Учитывая то, что в 3) было получено:

$$P[\underset{h\in B_i}{sup}|\Delta(h)| \ge \varepsilon] \le P[|\Delta(h_i)| \ge \tfrac{\varepsilon}{2}].$$

Тогда

$$P[\underset{h\in H}{sup}|\Delta(h)| \ge \varepsilon] \le 2\exp\left(\tfrac{-m\varepsilon^2}{2M^4}\right).$$