Answer to question 1

Answer to question 1a ($1\frac{1}{2}$ points for each subquestion).
The distribution function is (using the density from the Appendix):

$$F_{\mu,\sigma}(x) = \int_0^x (2\pi\sigma^2 y^2)^{-1/2} \exp\{-[\log(y) - \mu]^2/(2\sigma^2)\}dy = \int_0^x (2\pi\sigma^2 y^2)^{-1/2} \exp\{-[\log(y)]^2/(2\sigma^2)\}dy.$$

Apply the change-of-variables: $z = \sigma^{-1}\log(y)$. Thus, $dz = \sigma^{-1}y^{-1}dy$. Or, $dy = \sigma y dz$. This changes the integration domain from $(0, x)$ to $(-\infty, \sigma^{-1}\log(x))$. Put together:

$$\begin{aligned} F_{\mu,\sigma}(x) &= \int_0^x (2\pi\sigma^2 y^2)^{-1/2} \exp\{-[\log(y)]^2/(2\sigma^2)\}dy = \int_{-\infty}^{\sigma^{-1}\log(x)} (2\pi\sigma^2 y^2)^{-1/2} \exp(-z^2/2)\, \sigma\, y\, dz \\ &= \int_{-\infty}^{\sigma^{-1}\log(x)} (2\pi)^{-1/2} \exp(-z^2/2)\, dz = \Phi_{0,1}[\sigma^{-1}\log(x)]. \end{aligned}$$

This distribution function is continuous and monotone in x *($\frac{1}{2}$ point)*. Hence, the quantile function is then the inverse of $F_{\mu,\sigma}(x)$:

$$\begin{aligned} \alpha &= F_{\mu,\sigma}(x_\alpha) \\ \Leftrightarrow\ \alpha &= \Phi_{0,1}[\sigma^{-1}\log(x_\alpha)] \\ \Leftrightarrow\ \Phi_{0,1}^{-1}(\alpha) &= \sigma^{-1}\log(x_\alpha) \\ \Leftrightarrow\ \exp[\sigma\Phi_{0,1}^{-1}(\alpha)] &= x_\alpha. \end{aligned}$$

Thus, the quantile function is $F^{-1}(\alpha) = \exp[\sigma\Phi_{0,1}^{-1}(\alpha)]$.

Answer to question 1b
The first order (population) moment of the lognormal distributed X_i (from the Appendix) is $\mathbb{E}(X) = \exp(\mu^2 + \frac{1}{2}\sigma^2)$. The first order sample moment is the sample mean $\bar{X} = \frac{1}{n}\sum_{i=1} X_i$. Equate the population and sample moment, solve for σ^2 and obtain: $\hat{\sigma}^2_{MoM} = 2\log(\bar{X})$.

Answer to question 1c
Denote the inverse gamma prior on σ^2 by $\pi(\sigma^2)$. The Bayes estimator then is:

$$\mathbb{E}(\sigma^2 | X_1 = x_1, \ldots, X_n = x_n) = \int_0^\infty \sigma^2 \frac{\pi(\sigma^2)\ P(X_1 = x, \ldots, X_n = x_n \,|\, \sigma^2)}{\int_0^\infty \pi(\sigma^2)\ P(X_1 = x, \ldots, X_n = x_n \,|\, \sigma^2)\ d\sigma^2}\ d\sigma^2.$$

The denominator is:

$$\begin{aligned} &\int_0^\infty \pi(\sigma^2)\ P(X_1 = x, \ldots, X_n = x_n \,|\, \sigma^2)\ d\sigma^2 \\ &= \int_0^\infty \pi(\sigma^2) \prod_{i=1}^n P_{\sigma^2}(X_i = x_i \,|\, \sigma^2)\ d\sigma^2 \\ &= \int_0^\infty \beta^\alpha[\Gamma(\alpha)]^{-1}(\sigma^2)^{-\alpha-1}\exp(-\beta/\sigma^2) \prod_{i=1}^n (2\pi\sigma^2 x_i^2)^{-1/2} \exp\{-[\log(x_i)]^2/(2\sigma^2)\}\ d\sigma^2 \\ &= \beta^\alpha[\Gamma(\alpha)]^{-1}\Big[\prod_{i=1}^n (2\pi x_i^2)^{-1/2}\Big] \int_0^\infty (\sigma^2)^{-\alpha-1-n/2}\ \exp\{-\sigma^{-2}(\beta + \frac{1}{2}\sum_{i=1}^n [\log(x_i)]^2)\}\ d\sigma^2 \\ &= \beta^\alpha[\Gamma(\alpha)]^{-1}\Big[\prod_{i=1}^n (2\pi x_i^2)^{-1/2}\Big]\Big\{\beta + \frac{1}{2}\sum_{i=1}^n [\log(x_i)]^2\Big\}^{-\alpha-n/2} \Gamma(\alpha + n/2) \\ &\quad \times \int_0^\infty \Big\{\beta + \frac{1}{2}\sum_{i=1}^n [\log(x_i)]^2\Big\}^{\alpha+n/2} [\Gamma(\alpha + n/2)]^{-1}(\sigma^2)^{-\alpha-1-n/2}\ \exp\{-\sigma^{-2}(\beta + \frac{1}{2}\sum_{i=1}^n [\log(x_i)]^2)\}\ d\sigma^2 \\ &= \beta^\alpha[\Gamma(\alpha)]^{-1}\Big[\prod_{i=1}^n (2\pi x_i^2)^{-1/2}\Big]\Big\{\beta + \frac{1}{2}\sum_{i=1}^n [\log(x_i)]^2\Big\}^{-\alpha-n/2} \Gamma(\alpha + n/2), \end{aligned}$$

where in the last step the integral vanishes as the integrand is an inverse gamma density. By the same token the numerator is:

$$
\begin{aligned}
\int_0^\infty \sigma^2 \ \pi(\sigma^2) \prod_{i=1}^n P_{\sigma^2}(X_i = x_i \,|\, \sigma^2) \ d\sigma^2 & \\
= \ \beta^\alpha [\Gamma(\alpha)]^{-1} \Big[\prod_{i=1}^n (2\pi x_i^2)^{-1/2} \Big] \Big\{ \beta + \frac{1}{2} \sum_{i=1}^n [\log(x_i)]^2 \Big\}^{-\alpha - n/2 + 1} \Gamma(\alpha + n/2 - 1).
\end{aligned}
$$

The Bayes estimator is then given by division of the numerator by the denominator:

$$
\hat{\sigma^2}_B \ = \ \Big\{ \beta + \frac{1}{2} \sum_{i=1}^n [\log(x_i)]^2 \Big\} \Gamma(\alpha + n/2 - 1) [\Gamma(\alpha + n/2)]^{-1} = \Big\{ \beta + \frac{1}{2} \sum_{i=1}^n [\log(x_i)]^2 \Big\} (\alpha + n/2 - 1)^{-1},
$$

in which the property of the Gamma-function has been used.

Note: The argument above may be abbreviated, as the normalization constant does not affect the shape of the posterior.

Answer to question 2

Answer to question 2a

Statistical model: $X_i \sim \mathcal{N}(\mu, \sigma^2)$ and $Y_i \sim \mathcal{N}(\nu, \sigma^2)$ with unknown $\mu, \nu \in \mathbb{R}$ and unkown $\sigma > 0$. Assumption (!): variances are equal. Hypothesis: $H_0 : \mu = \nu$ and $H_a : \mu \neq \nu$.

Answer to question 2b

Remark 1: should a question 2a state a one-sided hypothesis, proceed one-sided.

Remark 2: One point for parts i), ii), iv), and two points for iii).

The test statistic is $T = (\bar{X} - \bar{Y})/(s_{x,y}\sqrt{n^{-1} + m^{-1}})$, which (under H_0) follows a t_{n+m-2} distribution. As $s_{x,y}^2 = (196.56 + 160.22)/16 = 22.24$, $s_{x,y} = 4.72$. The observed test statistic $T = t = (45.22 - 41.56)/(4.72\sqrt{\frac{1}{9} + \frac{1}{9}}) = 1.65$. On the other hand, the critical region is: $K_T = (-\infty, t_{16;0.025}] \cup [t_{16;0.975}, \infty) = (-\infty, -2.12] \cup [2.12, \infty)$. Thus, $T \notin K_T$ and H_0 is not rejected.

Answer to question 2c

The power function of a test with test statistic T and rejection region K_T is $\theta \mapsto \pi(\theta; K_T) = P_\theta(T \in K_T)$. The power at a $\theta \in \Theta_{H_a}$ can be increased by increasing the sample size: include more athletes! *(One point for definition, one for sample size argument).*

Answer to question 3

Answer to question 3a

First derive the maximum likelihood estimator *(Three points)*. The likelihood, its logarithm and the 1^{st} and 2^{nd} derivative of the latter are:

$$
\begin{aligned}
L(\theta; X_1, \ldots, X_n) &= \prod_{i=1}^n \ (2\pi\theta)^{-1/2} \exp[-x_i^2/(2\theta)], \\
\log[L(\theta; X_1, \ldots, X_n)] &\propto -\frac{n}{2}\log(\theta) - \frac{1}{2\theta}\sum_{i=1}^n x_i^2, \\
\frac{\partial}{\partial\theta}\log[L(\theta; X_1, \ldots, X_n)] &\propto -\frac{n}{2\theta} + \frac{1}{2\theta^2}\sum_{i=1}^n x_i^2, \\
\frac{\partial^2}{\partial\theta^2}\log[L(\theta; X_1, \ldots, X_n)] &\propto \frac{n}{2\theta^2} - \frac{1}{\theta^3}\sum_{i=1}^n x_i^2.
\end{aligned}
$$

Equating the derivate to zero and solving for θ yields: $\hat{\theta}_{ML} = \frac{1}{n}\sum_{i=1}^{n} x_i^2$. To verify this indeed corresponds to the location of a maximum of the (log-)likelihood evaluate the 2^{nd} order derivative at $\hat{\theta}_{ML}$, which is (indeed) smaller than zero.
The likelihood ratio statistic *(Two points)* is:

$$\begin{aligned}\lambda_n(X_1,\ldots,X_n) &= \frac{L(\hat{\theta}_{ML};X_1,\ldots,X_n)}{L(1;X_1,\ldots,X_n)} = \frac{\prod_{i=1}^{n} \ (2\pi\hat{\theta}_{ML})^{-1/2}\exp[-x_i^2/(2\hat{\theta}_{ML})]}{\prod_{i=1}^{n} \ (2\pi)^{-1/2}\exp(-x_i^2/2)}. \\ &= (\hat{\theta}_{ML})^{-n/2}\exp\Big(-\frac{1-\hat{\theta}_{ML}}{2\hat{\theta}_{ML}}\sum_{i=1}^{n} x_i^2\Big).\end{aligned}$$

Answer to question 3b
From lecture theorem: $2\log[\lambda_n(X_1,\ldots,X_n)] \rightsquigarrow \chi_1^2$ if $n\to\infty$ as $\Theta=(0,\infty)$ and $\{1\}$ is an interior point of Θ (i.e. the H_0 restriction is not on the boundary). For testing $H_0:\theta=1$ vs. $H_a:\theta\neq 1$ the likelihood-ratio test rejects if $2\log[\lambda_n(X_1,\ldots,X_n)] \geq \chi^2_{1,1-\alpha}$. The theorem on the correspondence between tests and confidence intervals then yields the confidence interval for θ:

$$\{\theta : \log[L(\theta;X_1,\ldots,X_n)] - \log[L(\hat{\theta}_{ML};X_1,\ldots,X_n)] \geq -\tfrac{1}{2}\chi^2_{1,1-\alpha}\}.$$

Answer to question 3c
Let random variable X follow p_θ. A pivot is a function $(X,\theta)\mapsto T(X,\theta)$ such that the probability distribution of $T(X,\theta)$ does not depend on θ or other unknown parameters. A pivotal quantity $T(X,\theta)$ is used to construct an exact (!) confidence interval for θ. When a pivot does not exist, an approximate confidence interval is constructed on the basis of an approximate pivot. For instance, for large samples the distribution of the pivot may be approximated by its asymptotic distribution (which ought not to depend on the unknown parameters). Hence, pivots yield exact confidence intervals, whereas approximate pivots produce approximate confidence intervals. As the confidence interval here is based on an asymptotic approximation: an approximate pivot.

Answer to question 4
Answer to question 4a
First show the density of the data (likelihood) belongs to the exponential family:

$$\begin{aligned}p_\theta(X_1,\ldots,X_n) &= \prod_{i=1}^{n}\theta^{-1}\exp(-x_i/\theta) = \theta^{-n}\exp(-\theta^{-1}\sum_{i=1}^{n} x_i) \\ &= c(\theta)\exp[Q_1(\theta)V_1(X_1,\ldots,X_n)],\end{aligned}$$

with $Q_1(\theta)=-\theta^{-1}$ and $V_1(X_1,\ldots,X_n)=\sum_{i=1}^{n} x_i$. As $\{-\theta^{-1}:\theta>0\}$ contains an interior point in $\mathbb{R}$, the statistic $\sum_{i=1}^{n} x_i$ is (by theorem from lecture) sufficient and complete.

Answer to question 4b
The maximum likelihood estimator $\hat{\theta}_{ML}=\overline{X}$ is based on $V_1(X_1,\ldots,X_n)=\sum_{i=1}^{n} x_i$, which is sufficient and complete. If $\hat{\theta}_{ML}$ is unbiased, then (by theorem from lecture) $\hat{\theta}_{ML}$ is a UMVU estimator. Note: $\mathbb{E}(\hat{\theta}_{ML})=\mathbb{E}(\overline{X})=\mathbb{E}(X_1)=\theta$. Thus, $\hat{\theta}_{ML}=\overline{X}$ is a UMVU estimator.

Answer to question 4c
Given $i_\theta=\mathbb{V}_\theta[\dot{\ell}_\theta(X_1)]=-\mathbb{E}[\ddot{\ell}_\theta(X_1)]$. The likelihood is:

$$\ell_\theta(X_1) = \log[\theta^{-1}\exp(-X_1/\theta)] = -\log(\theta)-X_1/\theta.$$

Its gradient is then $\dot{\ell}_\theta(X_1)=-\theta^{-1}+X_1\theta^{-2}$, while its second order derivative is $\ddot{\ell}_\theta(X_1)=\theta^{-2}-2X_1\theta^{-3}$. The Fisher information is then:

$$i_\theta = -\mathbb{E}[\ddot{\ell}_\theta(X_1)] = -\theta^{-2}+2\mathbb{E}(X_1)\theta^{-3} = -\theta^{-2}+2\theta\theta^{-3} = \theta^{-2}.$$

Answer to question 4d
The variance of the ML-estimator is $\mathbb{V}_\theta(\overline{X}) = \frac{1}{n}\mathbb{V}_\theta(X_1) = \theta^2/n$ (if $\exp(\theta^{-1})$ exists). From the lemma we have: $\theta^{-2} = \mathbb{V}(-\theta^{-1} + X_1\theta^{-2}) = \theta^{-4}\mathbb{V}(X_1)$. Thus, $\mathbb{V}(X_1) = \theta^2$.

Answer to question 4e
Cramer-Rao tells us that $\mathbb{V}_\theta(T) \geq (n\ i_\theta)^{-1}$ for unbiased estimators T of θ. Here the underbound is $(n\ i_\theta)^{-1} = \theta^2/n = \mathbb{V}(\overline{X}) = \mathbb{V}(\hat{\theta}_{ML})$. Hence, the underbound is sharp.