

10th problem set for Probability and Statistics — May 4/6

Summary

- **Method of Moments Estimation** solves the equation $m_1(\theta) = \widehat{m}_1(\theta)$ for the unknown θ .
 - Or a system of equations $m_r(\theta) = \widehat{m}_r(\theta)$ for $r = 1, 2, \dots$ as needed.
 - $L(\theta; x_1, \dots, x_n) = P(X_1 = x_1 \& \dots \& X_n = x_n)$... likelihood of observed data dependent of parameter θ .
 - or $L(\dots) = f_{X_1, \dots, X_n}(x_1, \dots, x_n)$... probability density function ...
 - $\ell(\theta; x_1, \dots, x_n) = \log L(\dots)$... for easier computations.
 - **Maximum Likelihood Estimation (MLE)** searches for θ for which $L(\theta; x_1, \dots, x_n)$, or $\ell(\dots)$, is maximized. Usually done using derivatives of L , or ℓ .
 - **Bias:** $\mathbb{E}_\theta(\hat{\theta} - \theta)$... θ true parameter, $\hat{\theta}$ our estimate (random variable as it depends on observed data).
 - **MSE (Mean Square Error):** $\mathbb{E}_\theta((\hat{\theta} - \theta)^2)$
 - Theorem: $MSE = bias(\hat{\theta})^2 + \text{var}(\hat{\theta})$.
-

1. We have a random sample $X_1, \dots, X_n \sim U(0, \vartheta)$.
 - (a) Propose a point estimate ϑ using the method of moments.
 - (b) Propose a point estimate ϑ using the method of maximum likelihood.
 - (c) For each of them, determine whether it is unbiased and consistent.
 - (d) Compute the Mean Square Error (MSE) for each of them.
 - (e) Which estimate is better? Can you think of an even better one?
2. For a random sample $X_1, \dots, X_n \sim Geom(p)$.
 - (a) Propose a point estimate ϑ using the method of moments.
 - (b) Propose a point estimate ϑ using the method of maximum likelihood.
 - (c) For each of them, determine whether it is unbiased and consistent.
3. For a random sample $X_1, \dots, X_n \sim Exp(1/\theta)$.
 - (a) Propose a point estimate θ using the method of moments.
 - (b) Propose a point estimate θ using the method of maximum likelihood.
 - (c) For each of them, determine whether it is unbiased and consistent.
 - (d) Compute the Mean Square Error (MSE).
4. We have a random sample $X_1, \dots, X_n \sim Exp(\lambda)$. We are interested in the probability p that $X > 1$ for $X \sim Exp(\lambda)$. (Recall that $p = e^{-\lambda \cdot 1}$.)
 - (a) Propose a point estimate for p (using any method), or several estimates if desired.
 - (b) Investigate its properties.
5. Let $S = \sum_{k=0}^{30} \binom{100}{k}$. Also, let $X = \sum_{i=1}^{100} X_i$, where X_i is 0 or 1, both with probability 1/2 and the variables X_1, \dots, X_n are independent. Thus, $X \sim Bin(100, 1/2)$.
 - (a) Express S using the cumulative distribution function F_X .
 - (b) Use CLT to estimate this probability.
 - (c) Alternatively, compute S using appropriate software and compare.

Additional Exercises

6. Let $X \sim \text{Exp}(\lambda)$ describe the distance traveled by a radioactive particle before decay. Our device captures its decay (and the decay location, i.e., the value of X), but only if $1 \leq X \leq 2$. Formally, we will consider a random sample $X_1, \dots, X_n \sim F_{X|B}$ for the event $B = 1 \leq X \leq 2$.

- Propose a point estimate λ using the method of moments.
- Propose a point estimate λ using the method of maximum likelihood.
- For each of them, determine whether it is unbiased and consistent.

7. We have a random sample $X_1, \dots, X_n \sim \text{Pois}(\lambda)$.

- Propose a point estimate for λ using the method of moments.
- Propose a point estimate for λ using the method of maximum likelihood.
- Calculate the mean squared error (MSE).

Interval estimates (NOT for the tutorial test)

8. We have one measurement $X \sim N(\mu, 1)$. (That is, the parameter $\vartheta = \mu$.)

- Find a 95% confidence interval for μ . (For specificity: we measured $x = 2.9$.)
- Instead of one measurement, we perform n (of course, independent) measurements. What will be the confidence interval for μ now? For specificity: we measured $x_1, \dots, x_9 = 1.82, 1.00, 2.50, 3.00, 0.50, 2.97, 1.76, 1.35, 3.41$.
- Let X still have mean μ and variance 1, but it is no longer necessarily normal. What changes?

9. This time we are sampling from $N(\mu, \sigma^2)$: we do not know μ or σ , so the parameter is $\vartheta = (\mu, \sigma)$. We measured values 8.47, 10.91, 10.87, 9.46, 10.40.

- Calculate the sample mean and sample variance.
- If we believed that the computed sample variance is the true value of σ^2 , find a confidence interval for μ .
- Find a confidence interval for μ using the Student's t -distribution.

10. We model the number of emails per day using a Poisson distribution $\text{Pois}(\lambda)$. In the first week of May, we received 34, 35, 29, 31, 30 emails. Find a 95% confidence interval for λ .

Use the last method from the lecture – the one using the Student's distribution. (Although the Poisson distribution is not normal, for sufficiently high values of λ , it is very similar to normal, so the method will have reliability close to 95%.)

p	0.9	0.95	0.975	0.99	0.995
$\Phi^{-1}(p)$	1.28	1.64	1.96	2.33	2.58
$\Psi_1^{-1}(p)$	3.08	6.31	12.71	31.82	63.66
$\Psi_2^{-1}(p)$	1.89	2.92	4.3	6.96	9.92
$\Psi_4^{-1}(p)$	1.53	2.13	2.78	3.75	4.6
$\Psi_8^{-1}(p)$	1.4	1.86	2.31	2.9	3.36
$\Psi_{20}^{-1}(p)$	1.33	1.72	2.09	2.53	2.85
$\Psi_{30}^{-1}(p)$	1.31	1.7	2.04	2.46	2.75