Statistics 630 – Final Exam
Monday, 13 December 2021

Printed Name: ________________________ **Email:** ________________________

INSTRUCTIONS FOR THE STUDENT:

1. You have 2 hours to complete the exam.
2. There are 10 pages including this cover sheet and the formula sheets.
3. Questions 1–8 are multiple choice and worth 5 points each.
4. Questions 9 (3 parts), 10 and 11 require solutions to be worked out, and are 10 points per part. (90 total points for the exam.)
5. Please write out your answers in *the spaces provided*, explaining your steps. You may refer to theorems by name/description rather than by its number in the book.
6. If you *cannot* print out the exam, please write your answers on blank sheet of paper – in order.
7. You may use the *attached formula sheets.* No other resources are allowed. Do not use the textbook, the class notes, homework or formula sheets that were posted online.
8. You may use but mostly do not need a calculator. You may leave answers in forms that can easily be put into a calculator such as $\frac{12}{19}$, $\binom{40}{5}$, e^{-3}, $\Phi(1.5)$, etc.
9. Do not discuss or provide any information to anyone concerning any of the questions on this exam until your solutions are returned or I post my solutions.

Questions 1–8 are multiple choice: circle the single correct answer. No partial credit!

1. (5 points) Suppose $X_1, \ldots, X_n$ is a random sample from a distribution such that $\mathsf{E}(X_i^2) = \theta^2$. Based on this, one example of a method of moments estimator for θ is
   (a) $\overline{X}$.
   (b) $\overline{X}^2$.
   (c) $\frac{1}{n}\sum_{i=1}^n X_i^2$.
   (d) $\left(\frac{1}{n}\sum_{i=1}^n X_i^2\right)^{1/2}$.
   (e) $\left(\frac{1}{n}\sum_{i=1}^n X_i^2\right)^2$.

2. (5 points) Consider a hypothesis test for $H_0 : \gamma = \gamma_0$ versus $H_a : \gamma > \gamma_0$. Everything else being equal, if the size α of the test is increased then the power function for the test will
   (a) stay the same as before.
   (b) increase.
   (c) decrease.
   (d) increase only for values of $\gamma < \gamma_0$.
   (e) decrease only for values of $\gamma > \gamma_0$.

3. (5 points) 10,000 rubber ducks are floated down a stream. 200 of the ducks are gold colored; the rest are yellow. Assuming they get thoroughly mixed during their passage, the number of gold ducks among the first 100 to cross the finish line has approximately
   (a) Bernoulli(0.02) distribution.
   (b) binomial(10000, 0.02) distribution.
   (c) normal(2, 1.96) distribution.
   (d) Poisson(2) distribution.
   (e) geometric(0.02) distribution.

4. (5 points) A simple random sample $V_1, \ldots, V_{80}$ of 80 exponential(λ) data has mean 4.35 and variance 15.89. As we have seen, the MLE for the distribution mean, $\mu = \frac{1}{\lambda}$, is $\widehat{\mu} = \overline{V}$. The value of the Wald 95% confidence interval for μ is
   (a) $4.35 \pm 1.960\sqrt{\frac{4.35}{80}}$.
   (b) $4.35 \pm 1.960\sqrt{\frac{15.89}{80}}$.
   (c) $4.35 \pm 1.960\frac{4.35}{\sqrt{80}}$.
   (d) $\frac{1}{4.35} \pm 1.960\frac{4.35}{\sqrt{80\times 15.89}}$.
   (e) $\frac{1}{4.35} \pm 1.960\frac{1}{4.35\sqrt{80}}$.

5. (5 points) Estimator $\widetilde{\theta}_1$ has bias $\frac{2}{n}$ and estimator $\widetilde{\theta}_2$ has bias $-\frac{1}{n}$. We create a new, unbiased, estimator $\widetilde{\theta}_3 = \frac{1}{3}\widetilde{\theta}_1 + \frac{2}{3}\widetilde{\theta}_2$. This estimator has MSE (mean squared error)

(a) $\frac{1}{3}\,\mathsf{Var}(\widetilde{\theta}_1) + \frac{2}{3}\,\mathsf{Var}(\widetilde{\theta}_2)$.

(b) $\frac{1}{9}\,\mathsf{Var}(\widetilde{\theta}_1) + \frac{4}{9}\,\mathsf{Var}(\widetilde{\theta}_2)$.

(c) $\frac{1}{3}\,\mathsf{Var}(\widetilde{\theta}_1) + \frac{2}{3}\,\mathsf{Var}(\widetilde{\theta}_2) + \frac{2}{3}\,\mathsf{Cov}(\widetilde{\theta}_1, \widetilde{\theta}_2)$.

(d) $\frac{1}{9}\,\mathsf{Var}(\widetilde{\theta}_1) + \frac{4}{9}\,\mathsf{Var}(\widetilde{\theta}_2) + \frac{2}{9}\,\mathsf{Cov}(\widetilde{\theta}_1, \widetilde{\theta}_2)$.

(e) $\frac{1}{9}\,\mathsf{Var}(\widetilde{\theta}_1) + \frac{4}{9}\,\mathsf{Var}(\widetilde{\theta}_2) + \frac{4}{9}\,\mathsf{Cov}(\widetilde{\theta}_1, \widetilde{\theta}_2)$.

6. (5 points) $T_1, \ldots, T_n$ is a random sample from a distribution with pdf $f(t) = \frac{\beta}{t^2} e^{-\beta/t}$, $t > 0$, $\beta > 0$.

(a) $\sum_{i=1}^n T_i^{-1}$ is sufficient and $\beta \sum_{i=1}^n T_i^{-1}$ is a pivot.

(b) $\sum_{i=1}^n T_i^{-1}$ is a pivot and $\beta \sum_{i=1}^n T_i^{-1}$ is sufficient.

(c) $1/\sum_{i=1}^n T_i$ is sufficient and $\beta/\sum_{i=1}^n T_i$ is a pivot.

(d) $\sum_{i=1}^n T_i$ is a pivot and $\frac{1}{\beta}\sum_{i=1}^n T_i$ is sufficient.

(e) $\sum_{i=1}^n T_i$ is sufficient and $\beta \sum_{i=1}^n T_i^{-1}$ is a pivot.

7. (5 points) $V_1, \ldots, V_n$ is a random sample from a distribution with mean $1/(5\theta)$, and has score function $S(\theta) = n\theta - 5n\theta^2\overline{V}$ and Fisher information $I_n(\theta) = n$. The size 0.01 score test for $H_0 : \theta = 2$ versus $H_2 : \theta \neq 2$ has rejection criterion

(a) $\sqrt{n}(20\overline{V} - 2) > z_{0.99}$.

(b) $\sqrt{n}(20\overline{V} - 2) > z_{0.98}$.

(c) $n(20\overline{V} - 2)^2 > \chi^2_{0.99}(1)$.

(d) $n(5\theta^2\overline{V} - \theta)^2 > \chi^2_{0.99}(n)$.

(e) $\sqrt{n}(5\theta^2\overline{V} - \theta)^2 > \chi^2_{0.99}(n)$.

8. (5 points) The conditional distribution of X, given $Y = y$, is Poisson$(y + y^2)$ and Y has some distribution F_Y. Suppose one only observes a random sample $Y_1, \ldots, Y_n$ from F_Y. Then an unbiased estimator for $\mathsf{E}(X)$ (<u>not</u> $\mathsf{E}(Y)$) is

(a) $\overline{X}$.

(b) $\overline{Y}$.

(c) $\frac{1}{n-1}\sum_{i=1}^n (X_i - \overline{X})^2$.

(d) $\overline{Y} + \overline{Y}^2$.

(e) $\frac{1}{n}\sum_{i=1}^n (Y_i + Y_i^2)$.

Provide solutions to Questions 9–11, to the point of a calculable expression.

9. Suppose $W_1, \ldots, W_n$ are a random sample from distribution with pdf $f(w) = \frac{4}{\sqrt{\pi}}\beta^{-3}w^2e^{-(w/\beta)^2}$, $w > 0$, $\beta > 0$. Note: $2W_i^2/\beta^2 \sim$ chi-square(3).

   (a) (10 points) The MLE for β is $\widehat{\beta} = \left(\frac{2}{3n}\sum_{i=1}^n W_i^2\right)^{1/2}$. Is $\widehat{\beta}^2$ unbiased for β^2? Is $\widehat{\beta}$ unbiased for β? Explain fully. (Use the note above.)

   (b) (10 points) The log-likelihood function is

$$\ell(\beta) = \sum_{i=1}^n \log(4W_i^2/\sqrt{\pi}) - 3n\log\beta - \frac{1}{\beta^2}\sum_{i=1}^n W_i^2.$$

   Find the Fisher information $I_n(\beta)$ for the sample.

(9. continued)

(c) (10 points) Now determine the form of the size α generalized likelihood ratio test for $H_0 : \beta = \beta_0$ versus $H_a : \beta \neq \beta_0$. (Express in terms of $\widehat{\beta}$ and β_0, simplified.)

10. (10 points) $X_1, \ldots, X_n$ are iid Poisson(λ) random variables. Recall that $Y = \sum_{i=1}^{n} X_i$ is sufficient for λ. Assume λ has an exponential(1) *prior* distribution. Show that λ has a gamma *posterior* distribution. What are the posterior parameters and the posterior mean?

(One more problem next page)

11. (10 points) Let (S,T) be a random pair with joint pdf $f(s,t) = \frac{s+t}{2}e^{-s-t}$, $s > 0$, $t > 0$. Find $\mathsf{E}(S+T)$.

Formulas for Final Exam

Bayes' rule $\mathsf{P}(B_j \mid A) = \frac{\mathsf{P}(A|B_j)\mathsf{P}(B_j)}{\sum_{k=1}^{n} \mathsf{P}(A|B_k)\mathsf{P}(B_k)}$ if $B_1, \ldots, B_n$ are disjoint and $\bigcup_{k=1}^{n} B_k = S$.

quantile function $Q_X(p)$ satisfies $F_X(x) \leq p \leq F(Q_X(p))$ for all $x < Q_X(p)$. $F(Q_X(p)) = p$ if X is a continuous rv.

distribution of a function of X $F_Y(y) = \mathsf{P}(h(X) \leq y)$ for $Y = h(X)$.

If X is a discrete rv or $h(x)$ takes only countably many values then Y has pmf $p_Y(y) = \mathsf{P}(h(X) = y)$.

If X is a continuous rv and $h(x)$ is a continuous function then Y has pdf $f_Y(y) = \frac{\mathrm{d}x}{\mathrm{d}y}\,\mathsf{P}(h(X) \leq y)$.

binomial theorem $\sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k} = (a+b)^n$.

geometric sum $\sum_{k=n}^{\infty} a^k = \frac{a^n}{1-a}$ if $-1 < a < 1$.

exponential expansion $\sum_{k=0}^{\infty} \frac{a^k}{k!} = \mathrm{e}^a$.

gamma integral $\int_0^\infty x^a \mathrm{e}^{-x}\,\mathrm{d}x = \Gamma(a+1) = a!$ for $a > -1$.

Bernoulli pmf $p(x) = (1-\theta)^{1-x}\theta^x I_{\{0,1\}}(x)$ for $0 < \theta < 1$, same as binomial$(1,\theta)$.

beta(a,b) pdf $f(x) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} x^{a-1}(1-x)^{b-1} I_{(0,1)}(x)$ for $a > 0$, $b > 0$; $\mathsf{E}(X) = \frac{a}{a+b}$ $\mathsf{Var}(X) = \frac{ab}{(a+b)^2(a+b+1)}$.

binomial(n, θ) pmf $p(x) = \binom{n}{x}\theta^x(1-\theta)^{n-x}I_{\{0,1,\ldots,n\}}(x)$ for $0 < \theta < 1$. $\mathsf{E}(X) = n\theta$, $\mathsf{Var}(X) = n\theta(1-\theta)$, $m(s) = (1-\theta+\theta\mathrm{e}^s)^n$.

chi-square(n) same as gamma($\frac{n}{2}, \frac{1}{2}$), the distribution of $X = Z_1^2 + \cdots + Z_n^2$ for iid standard normal $Z_1, \ldots, Z_n$. $\mathsf{E}(X) = n$, $\mathsf{Var}(X) = 2n$.

In particular, if $X_1, \ldots, X_n \overset{\text{iid}}{\sim} \text{normal}(\mu, \sigma^2)$ then $\frac{(n-1)S^2}{\sigma^2} \sim \text{chi-square}(n-1)$.

discrete uniform(N) pmf $p(x) = \frac{1}{N}\, I_{\{1,2,\ldots,N\}}(x)$. $\mathsf{E}(X) = \frac{N+1}{2}$, $\mathsf{Var}(X) = \frac{N^2-1}{12}$.

exponential(λ) pdf $f(x) = \lambda\mathrm{e}^{-\lambda x}I_{(0,\infty)}(x)$ for $\lambda > 0$, same as gamma($1, \lambda$). $\mathsf{E}(X) = \frac{1}{\lambda}$, $\mathsf{Var}(X) = \frac{1}{\lambda^2}$.

F(m, n) the distribution of $W = \frac{X/m}{Y/n}$ where $X \sim \text{chi-square}(m)$, $Y \sim \text{chi-square}(n)$, independent. $\mathsf{E}(W) = \frac{n}{n-2}$ if $n > 2$.

gamma(α, λ) pdf $f(x) = \frac{\lambda^\alpha}{\Gamma(\alpha)}\, x^{\alpha-1}\mathrm{e}^{-\lambda x}I_{(0,\infty)}(x)$ for $\lambda > 0$, $\alpha > 0$; $\mathsf{E}(X) = \frac{\alpha}{\lambda}$, $\mathsf{Var}(X) = \frac{\alpha}{\lambda^2}$, $m(s) = \left(\frac{\lambda}{\lambda - s}\right)^\alpha$ if $s < \lambda$.

geometric(θ) pmf $p(x) = \theta(1-\theta)^x I_{\{0,1,2,\ldots\}}(x)$ for $0 < \theta < 1$, same as negative binomial($1, \theta$). $\mathsf{E}(X) = \frac{1-\theta}{\theta}$, $\mathsf{Var}(X) = \frac{1-\theta}{\theta^2}$.

hypergeometric(N, M, n) pmf $p(x) = \frac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}}\, I_{\{0,1,\ldots,n\}}(x)$ for $M < N$. $\mathsf{E}(X) = np$ where $p = \frac{M}{N}$, $\mathsf{Var}(X) = \frac{N-n}{N-1}\, np(1-p)$.

negative binomial(r, θ) pmf $p(x) = \binom{r+x-1}{r-1}\theta^r(1-\theta)^x I_{\{0,1,2,\ldots\}}(x)$ for $0 < \theta < 1$. $\mathsf{E}(X) = \frac{r(1-\theta)}{\theta}$, $\mathsf{Var}(X) = \frac{r(1-\theta)}{\theta^2}$, $m(s) = \left(\frac{\theta}{1-(1-\theta)\mathrm{e}^s}\right)^r$ if $s < -\log(1-\theta)$.

normal(μ, σ^2) pdf $f(x) = \frac{1}{\sqrt{2\pi\sigma^2}}\, \mathrm{e}^{-(x-\mu)^2/(2\sigma^2)}I_{(-\infty,\infty)}(x)$ for $\sigma^2 > 0$; $\mathsf{E}(X) = \mu$, $\mathsf{Var}(X) = \sigma^2$, $m(s) = \mathrm{e}^{\mu s + \sigma^2 s^2/2}$.

Poisson(λ) pmf $p(x) = \frac{\lambda^x}{x!}\, \mathrm{e}^{-\lambda}I_{\{0,1,2,\ldots\}}(x)$ for $\lambda > 0$. $\mathsf{E}(X) = \lambda$, $\mathsf{Var}(X) = \lambda$, $m(s) = \mathrm{e}^{\lambda(\mathrm{e}^s - 1)}$.

t(n) the distribution of $T = \frac{Z}{\sqrt{Y/n}}$ where $Z \sim \text{normal}(0,1)$, $Y \sim \text{chi-square}(n)$, independent. $\mathsf{E}(T) = 0$, $\mathsf{Var}(T) = \frac{n}{n-2}$ if $n > 2$. In particular, if $X_1, \ldots, X_n \overset{\text{iid}}{\sim} \text{normal}(\mu, \sigma^2)$ then $\frac{\bar{X}-\mu}{S/\sqrt{n}} \sim \mathrm{t}(n-1)$.

uniform(a, b) pdf $f(x) = \frac{1}{b-a}\, I_{(a,b)}(x)$ for $a < b$. $\mathsf{E}(X) = \frac{a+b}{2}$, $\mathsf{Var}(X) = \frac{(b-a)^2}{12}$.

Weibull(α, β) pdf $f(x) = \frac{\alpha}{\beta}\, (x/\beta)^{\alpha-1}\mathrm{e}^{-(x/\beta)^\alpha}I_{(0,\infty)}(x)$ for $\alpha > 0$, $\beta > 0$. $\mathsf{E}(X^k) = \beta^k\, \Gamma(1 + \frac{k}{\alpha})$.

marginal pmf/pdf $p_X(x) = \sum_y p(x,y)$; $f_X(x) = \int_{-\infty}^{\infty} f(x,y)\, dy$.

conditional pmf/pdf $p_{X|Y}(x|y) = \frac{p(x,y)}{p_Y(y)}$; $f_{X|Y}(x|y) = \frac{f(x,y)}{f_Y(y)}$.

independent random variables $p(x,y) = p_X(x)p_Y(y)$ if (X,Y) is discrete;
$f(x,y) = f_X(x)f_Y(y)$ if (X,Y) is continuous.

discrete convolution $p_{X+Y}(z) = \sum_x p_X(x)p_Y(z-x)$ for independent X, Y.

continuous convolution $f_{X+Y}(z) = \int_{-\infty}^{\infty} f_X(x)f_Y(z-x)\,dx$ for independent X, Y.

covariance and correlation $\mathsf{Cov}(X,Y) = \mathsf{E}((X-\mu_X)(Y-\mu_Y)) = \mathsf{E}(XY) - \mu_X\mu_Y$; $\mathsf{Corr}(X,Y) = \frac{\mathsf{Cov}(X,Y)}{\sigma_X\sigma_Y}$. For independent X and Y, $\mathsf{Cov}(X,Y) = \mathsf{Corr}(X,Y) = 0$.

expectation of a sum $\mathsf{E}(a_1X_1 + \cdots + a_nX_n) = a_1\mathsf{E}(X_1) + \cdots + a_n\mathsf{E}(X_n)$.

expectation of a product If $X_1, \ldots, X_n$ are independent, $\mathsf{E}\Big(\prod_{i=1}^n h_i(X_i)\Big) = \prod_{i=1}^n \mathsf{E}(h_i(X_i))$.

variance of a sum $\mathsf{Var}(aX + bY) = a^2\,\mathsf{Var}(X) + 2ab\,\mathsf{Cov}(X,Y) + b^2\,\mathsf{Var}(Y)$.

variance of a sum of independent rvs $\mathsf{Var}(a_1X_1 + \cdots + a_nX_n) = a_1^2\,\mathsf{Var}(X_1) + \cdots + a_n^2\,\mathsf{Var}(X_n)$.

moments k-th moment is $\mu_k = \mathsf{E}(X^k)$, $k = 1, 2, \ldots$

moment generating function $m_X(s) = \mathsf{E}(\mathrm{e}^{sX})$; $\mathsf{E}(X^k) = \frac{\mathrm{d}x^k}{\mathrm{d}s^k}\, m_X(s)\Big|_{s=0}$.

mgf of a sum If X and Y are independent, $m_{aX+bY}(s) = \mathsf{E}(\mathrm{e}^{(aX+bY)s}) = m_X(as)m_Y(bs)$.

conditional expectation $\mathsf{E}(h(Y)|X = x) = \sum_y h(y)p_{Y|X}(y|x)$ or $\mathsf{E}(h(Y) \mid X = x) = \int_{-\infty}^{\infty} h(y)f_{Y|X}(y|x)\,dy$.

iterated expectation $\mathsf{E}(h(Y)) = \mathsf{E}(\mathsf{E}(h(Y) \mid X))$, $\mathsf{E}(g(X)h(Y)) = \mathsf{E}(g(X)\mathsf{E}(h(Y) \mid X))$.

conditional variance $\mathsf{Var}(Y \mid X) = \mathsf{E}(Y^2 \mid X) - (\mathsf{E}(Y \mid X))^2$.

variance partition formula $\mathsf{Var}(Y) = \mathsf{E}(\mathsf{Var}(Y \mid X)) + \mathsf{Var}(\mathsf{E}(Y \mid X))$.

Markov's inequality $\mathsf{P}(|X| \geq x) \leq \frac{\mathsf{E}(|X|)}{x}$ for $x > 0$.

Chebyshev's inequality $\mathsf{P}(|X - \mu_X| \geq x) \leq \frac{\mathsf{Var}(X)}{x^2}$ for $x > 0$.

sample mean, variance, k-th moment $\bar{X} = \frac{1}{n}\sum_{i=1}^n X_i$; $\widehat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^n (X_i - \bar{X})^2$; $m_k = \frac{1}{n}\sum_{i=1}^n X_i^k$.

unbiased sample variance $S^2 = \frac{n}{n-1}\,\widehat{\sigma}^2 = \frac{1}{n-1}\sum_{i=1}^n (X_i - \bar{X})^2$.

law of large numbers For iid $X_1, X_2, \ldots$ with mean μ, $\bar{X}_n \to \mu$ as $n \to \infty$.

central limit theorem For iid $X_1, X_2, \ldots$ with mean μ and variance σ^2, $\mathsf{P}\Big(\frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}} \leq z\Big) = \mathsf{P}\Big(\frac{X_1 + \cdots + X_n - n\mu}{\sqrt{n}\sigma} \leq z\Big) \to \Phi(z)$ (normal(0,1) cdf), as $n \to \infty$.

bias and standard error $\mathsf{Bias}(\widehat{\theta}) = \mathsf{E}(\widehat{\theta}) - \theta$; $\mathsf{SE}(\widehat{\theta}) = \sqrt{\mathsf{Var}(\widehat{\theta})}$.

mean squared error $\mathsf{MSE}(\widehat{\theta}) = \mathsf{E}((\widehat{\theta} - \theta)^2) = \mathsf{Var}(\widehat{\theta}) + (\mathsf{Bias}(\widehat{\theta}))^2$.

consistency $\widehat{\theta}$ is consistent if $\mathsf{MSE}(\widehat{\theta}) \to 0$ as $n \to \infty$.

method of moments for iid sample match the k-th population moment $\mathsf{E}(X^k)$ with the k-th sample moment m_k, and solve for the desired parameter estimates.

likelihood function $L(\theta|X_1, \ldots, X_n) = \prod_{i=1}^n f_\theta(X_i)$ for iid sample $\underline{\mathrm{X}} = (X_1, \ldots, X_n)$.

maximum likelihood for iid sample maximize the likelihood function $L(\theta|X_1,\dots,X_n)=\prod_{i=1}^n f_\theta(X_i)$ or the log-likelihood $\ell(\theta|X_1,\dots,X_n)=\log L(\theta|X_1,\dots,X_n)=\sum_{i=1}^n \log f_\theta(X_i)$.

If $\log L(\theta)$ is differentiable and concave at θ, the MLE is a solution to $S(\theta)=\frac{d}{d\theta}\log L(\theta)=0$. (For a multidimensional parameter θ this is a system of equations.)

score function $S(\theta|X_1,\dots,X_n)=\frac{\mathrm{d}}{\mathrm{d}\theta}\,\ell(\theta)$.

Fisher information $I_n(\theta)=\mathsf{Var}(\frac{\mathrm{d}}{\mathrm{d}\theta}\,\ell(\theta))=-\mathsf{E}(\frac{\mathrm{d}^2}{\mathrm{d}\theta^2}\,\ell(\theta))$, if ℓ has two derivatives.

For an iid sample, $I_n(\theta)=nI_1(\theta)$ and $I_1(\theta)=\mathsf{Var}(\frac{\mathrm{d}}{\mathrm{d}\theta}\log f_\theta(X_1))=-\mathsf{E}(\frac{\mathrm{d}^2}{\mathrm{d}\theta^2}\log f_\theta(X_1))$.

sufficient statistic $T=T(\underline{X})$ is sufficient if $L(\theta|\underline{X})=h(\underline{X})g(T(\underline{X}),\theta)$ for some functions $h(\underline{x})$ and $g(t,\theta)$.

exponential family The pdf/pmf has the form $f_X(x|\theta)=d(\theta)h(x)\mathrm{e}^{c(\theta)t(x)}$ for all x,θ. In this case, with an iid random sample, $T(\underline{X})=\sum_{i=1}^n t(X_i)$ is a sufficient statistic and $I_n(\theta)=n(c'(\theta))^2\,\mathsf{Var}(t(X_1))$.

asymptotics for MLE Assuming Fisher information exists and $\widehat{\theta}$ is the MLE, $\widehat{\theta}\to\theta$ in probability and $\sqrt{I_n(\theta)}(\widehat{\theta}-\theta)\to$ normal$(0,1)$ in distribution as $n\to\infty$.

asymptotic normality $\widehat{\theta}$ is asymptotic normal(θ,V_n) if $\frac{\widehat{\theta}-\theta}{\sqrt{V_n}}\to$ normal$(0,1)$ in distribution as $n\to\infty$. V_n may depend on θ or other parameters. If $\widehat{V}_n$ is an estimator such that $\widehat{V}_n/V_n\to 1$ then $\frac{\widehat{\theta}-\theta}{\sqrt{\widehat{V}_n}}\to$ normal$(0,1)$ in distribution.

delta method If $g(\theta)$ is continuously differentiable and estimator $\widehat{\theta}$ is asymptotic normal(θ,V_n), then $g(\widehat{\theta})$ is asymptotic normal$(g(\theta),(g'(\theta))^2V_n)$.

level γ confidence interval $(L(\underline{X}),U(\underline{X}))$ such that $\mathsf{P}_\theta(L(\underline{X})\le\theta\le U(\underline{X}))=\gamma$.

confidence interval from pivot If $h(\underline{X},\theta)$ has a distribution that does not depend on θ, a level γ confidence interval is defined by $\{\theta : h(\underline{X},\theta)\in A\}$ where $\mathsf{P}_\theta(h(\underline{X},\theta)\in A)=\gamma$.

Wald confidence interval If $\widehat{\theta}$ is asymptotic normal(θ,V_n) and $\widehat{V}_n$ is an estimator for V_n, an approximate level γ confidence interval for θ has endpoints $\widehat{\theta}\pm z_{(1+\gamma)/2}\sqrt{\widehat{V}_n}$, where $z_{(1+\gamma)/2}$ is the $(1+\gamma)/2$ quantile of the normal(0,1) distribution.

score confidence interval For MLE $\widehat{\theta}$, an approximate level γ confidence interval defined by $\{\theta : -z_{(1+\gamma)/2}\le(\frac{\mathrm{d}}{\mathrm{d}\theta}\,\ell(\theta))/\sqrt{I_n(\theta)}\le z_{(1+\gamma)/2}\}$, where $z_{(1+\gamma)/2}$ is the $(1+\gamma)/2$ quantile of the normal(0,1) distribution.

A related method is the interval given by $\{\theta : -z_{(1+\gamma)/2}\le\sqrt{I_n(\theta)}(\widehat{\theta}-\theta)\le z_{(1+\gamma)/2}\}$.

Type I and II errors, level and power A Type I error is rejecting H_0 when it is true. The level of a test is $\alpha=\max_{\theta\in H_0}\mathsf{P}_\theta(H_0$ is rejected) computed with values of θ such that H_0 true.

A Type II error is not rejecting H_0 when H_a is true. The power of a test is $\beta=\beta(\theta)=\mathsf{P}_\theta(H_0$ is rejected) computed with parameter value θ (satisfying H_a).

***P*-value** The smallest level α for which H_0 will still be rejected – it is a statistic (function of the data).

Neyman-Pearson likelihood ratio test For simple hypotheses $H_0 : \theta = \theta_0$ vs. $H_a : \theta = \theta_1$, reject H_0 if $LR = \frac{L(\theta_1)}{L(\theta_0)} \geq c_\alpha$ where $\mathsf{P}(LR \geq c_\alpha) = \alpha$ when H_0 is true.
If, for each c, $LR \geq c \iff T \geq k$ (or $R \geq c \iff T \leq k$) for some statistic T and some value k then it suffices to find k_α such that $\mathsf{P}(T \geq k_\alpha) = \alpha$ (resp., $\mathsf{P}(T \leq k_\alpha) = \alpha$) when H_0 is true.

generalized likelihood ratio test For hypotheses H_0 and H_a about parameter θ and MLE $\widehat{\theta}$, reject H_0 if $LR = \frac{L(\widehat{\theta})}{\max_{\theta \in H_0} L(\theta)} \geq c_\alpha$ where $\max_{\theta \in H_0} \mathsf{P}(LR \geq c_\alpha) = \alpha$.
If $H_0 : \theta = \theta_0$ and $H_a : \theta \neq \theta_0$ then $LR = \frac{L(\widehat{\theta})}{L(\theta_0)}$.

uniformly most powerful test A test is UMP if it has maximum possible power for every parameter value θ that satisfies H_a.
In particular, if the test is the same as the Neyman-Pearson test for each θ satisfying H_a then it is UMP.

Wald test If $\widehat{\theta}$ is asymptotic normal(θ, V_n) and $\widehat{V}_n$ is an estimator for V_n, reject $H_0 : \theta = \theta_0$ when $\frac{|\widehat{\theta}-\theta_0|}{\sqrt{\widehat{V}_n}} \geq z_{1-\alpha/2}$, where $z_{1-\alpha/2}$ is the $(1-\alpha/2)$ quantile of the normal(0,1) dist. Equivalently, reject H_0 if $\frac{(\widehat{\theta}-\theta_0)^2}{\widehat{V}_n} \geq \chi^2_{1,1-\alpha}$.
Important case: $\widehat{\theta}$ is the MLE and $\widehat{V}_n = 1/I_n(\widehat{\theta})$.

(asymptotic) score test For MLE $\widehat{\theta}$, reject $H_0 : \theta = \theta_0$ when $\frac{|\frac{\mathrm{d}}{\mathrm{d}\theta}\,\ell(\theta_0)|}{\sqrt{I_n(\theta_0)}} \geq z_{1-\alpha/2}$, where $z_{1-\alpha/2}$ is the $(1-\alpha/2)$ quantile of the normal(0,1) dist. Equivalently, reject H_0 if $\frac{(\frac{\mathrm{d}}{\mathrm{d}\theta}\,\ell(\theta_0))^2}{I_n(\theta_0)} \geq \chi^2_{1,1-\alpha}$.
A related test is to reject $H_0 : \theta = \theta_0$ when $\sqrt{I_n(\theta_0)}\,|\widehat{\theta} - \theta_0| \geq z_{1-\alpha/2}$, and $\widehat{\theta}$ is the MLE.

asymptotic likelihood ratio test Using the generalized LR statistic, reject $H_0 : \theta = \theta_0$ when $2\log(LR) \geq z^2_{1-\alpha/2} = \chi^2_{1,1-\alpha}$.

test equivalent to interval Define a test from an interval (or an interval from a test) by: reject $H_0 : \theta = \theta_0$ at level $\alpha \iff \theta_0$ is not in the $1-\alpha$ confidence interval.

prior and posterior distributions If the prior density (or pmf) for θ is $f_\Theta(\theta)$ then the posterior density (or pmf) is $f_\Theta(\theta|\underline{\mathrm{X}}) = c(\underline{\mathrm{X}}) f_{\underline{\mathrm{X}}}(\underline{\mathrm{X}}|\theta) f_\Theta(\theta)$, with $c(\underline{\mathrm{X}})$ chosen so that $f_\Theta(\theta|\underline{\mathrm{X}})$ is a proper pdf (pmf) in θ.

Bayes estimator Either the mean or the mode of the posterior distribution.

Bayes γ credible interval An interval $(L(\underline{\mathrm{X}}), U(\underline{\mathrm{X}}))$ such that, under the posterior distribution, $\mathsf{P}(L(\underline{\mathrm{X}}) \leq \theta \leq U(\underline{\mathrm{X}}) \mid \underline{\mathrm{X}}) = \gamma$.
The interval is HPD (highest posterior density) if it equals the set $\{\theta : f_\Theta(\theta|\underline{\mathrm{X}}) \geq c\}$ for some constant c.

Bayes Hypothesis test Choose H_1 if and only if $\dfrac{\mathsf{P}(H_1 \mid \underline{\mathrm{X}})}{\mathsf{P}(H_0 \mid \underline{\mathrm{X}})} > 1$.