Homework #1 #1) $x_1, x_2 \sim f_\theta : f(x) = e^{-x} : x>0$ (a) follow Exp : $\lambda = 1$ (b1) $E(y) : y = x_1 - x_2 \Rightarrow E(x_1 - x_2) = E(x_1) - E(x_2) = \int x_1 e^{-x_1} dx_1 - \int x_2 e^{-x_2} dx_2 = 1 - 1 = 0$
(b2) $V(y) = V(x_1 - x_2) = V(x_1) + V(x_2) - 0 = 1+1 = 2$ (#2) $x_1, x_2, x_3 \sim$ Bernoulli$(p = \frac{1}{2})$ (a) $E(2x_1 - x_3 - 4) = \frac{-5}{2}$ (b) $E[(x_1 - 2x_2 + x_3)^2] =$
$E[x_1^2 + 4x_2^2 + x_3^2 - 4x_1x_2 + 2x_1x_3 - 4x_2x_3] = \frac{3}{2}$

Definitions $\Gamma(\alpha) = \int_0^\infty t^{\alpha-1} e^{t} dt$, $\Gamma(\alpha) = (\alpha - 1)!$

Final Review

#1) test null hyp. that device is effective (a) what conditions commit type I / type II errors? Ans:

	$H_0=T$	$H_0=F$
reject	I	✓
fail to reject	✓	II

(b) rephrase null so that Type I error becomes II Ans: H_0: "device isn't effective" *

#2) random sample assoc. w/ θ (a) definitions of unbiased estimator & asymptotically unbiased estimator? (a1) $E(\hat\theta)=\theta$ ... (b) Show $\frac{X+1}{n+2}$ is unbiased of θ, is asymptotically?
(b) $X\sim Binom(n,\theta)$, $E(X)=n\theta \Rightarrow E\left(\frac{X+1}{n+2}\right)=\frac{E(X)+1}{n+2}=\frac{n\theta+1}{n+2}=\frac{n}{n+2}\theta+\frac{1}{n+2}\neq\theta$ * ∴ biased estimator.
$\lim_{n\to\infty}E\left(\frac{X+1}{n+2}\right)=\frac{n}{n+2}\theta+\frac{1}{n+2}\xrightarrow{n\to\infty}\theta$ * since $\frac{n}{n+2}\to1$, $\frac{1}{n+2}\to0$

#3) 64 cacti, $\bar x=52.80$, $s=4.50$. 99% CI of the average growth (a) $n=64$, $\bar X=52.8$, $s=4.50$ σ unknown but $n\geq30$
by CLT: $\mu\in\bar x\pm z_{\frac{\alpha}{2}}\frac{s}{\sqrt n}\Rightarrow\mu\in52.8\pm2.575\frac{4.5}{\sqrt{64}}\Rightarrow\mu\in(51.352, 54.248)$ *

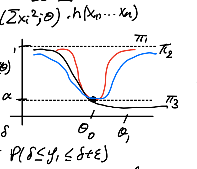
#4) sample mean size n, from Exp: $f(x)=\begin{cases}\frac{1}{\theta}e^{-\frac{x}{\theta}} & : x>0\\ 0 & : x\leq0\end{cases}$ is minimum unbiased estimator of θ. (a) $X\sim Exp(\theta)$, $E(X)=\theta$
1) $E(\bar X)=\theta$ ∴ unbiased, 2) show $V(\bar X)=\frac{1}{nE\left\{\left[\frac{\partial\log f(x;\theta)}{\partial\theta}\right]^2\right\}}$, CRLB $var(\bar X)=\frac{var(X)}{n}=\frac{\theta^2}{n}$, ⇒ $E\left\{\left[\frac{\partial\log f(x;\theta)}{\partial\theta}\right]^2\right\}=E\left[\frac{\partial}{\partial\theta}\left(\ln\left(\frac{1}{\theta}e^{-\frac{x}{\theta}}\right)\right)\right]^2$
$\Rightarrow E\left[\frac{\partial}{\partial\theta}\left(-\log\theta-\frac{x}{\theta}\right)\right]^2=E\left[-\frac{1}{\theta}+\frac{x}{\theta^2}\right]^2 \to \frac{1}{\theta^4}E(X-\theta)^2=\frac{Var(X)}{\theta^4}=\frac{\theta^2}{\theta^4}=\frac{1}{\theta^2}$
∴ CRLB is $\frac{1}{n\frac{1}{\theta^2}}=\frac{\theta^2}{n}=var(\bar X)$ *
∴ $\bar X$ is MVUE

#5) $H_0:\beta=\beta_0$, $H_1:\beta=\beta_0+2$. The null rejected iff r.v. $X>\beta_0+1$ find type I, II prob. $X\sim Unif[\alpha=0,\beta]$, $R=\{x>\beta_0+1\}$ (a) $P(\text{Type I error})=P(X>\beta_0|H_0)$
$=P(X>\beta_0+1|\beta=\beta_0)=0$ ∵ $X\sim Unif[0,\beta_0]$ under H_0 (b) $P(\text{Type II error})=P(X\leq\beta_0+1|H_1)=P(X\leq\beta_0+1|\beta=\beta_0+2)=\frac{\beta_0+1}{\beta_0+2}$ $|X\sim Unif[0,\beta_0+2]$

#6) $X_1\dots X_n\sim f_\theta$: $f(x)=\theta x^{\theta-1}$: $x\in(0,1)$, $\theta>0$. Find $\hat\theta_{MLE}$. (a) $L(\theta)=\prod_{i=1}^n(\theta x_i^{\theta-1})=\theta^n\left(\prod_{i=1}^n x_i\right)^{\theta-1}$
$\Rightarrow \ell(\theta)=\ln L(\theta)=n\ln(\theta)+(\theta-1)\ln\left(\prod_{i=1}^n x_i\right)\Rightarrow\frac{\partial\ell}{\partial\theta}=\frac{n}{\theta}+\sum\ln(x_i)=0\Rightarrow\theta=\frac{-n}{\sum\ln(x_i)}$
$\ell'(\theta)<0$: $\theta>\frac{-n}{\sum\ln(x_i)}$, $\ell'(\theta)\geq0$: $\theta\leq\frac{-n}{\sum\ln(x_i)}$ ∴ $\theta=\frac{-n}{\sum\ln(x_i)}$ = global max $=\hat\theta_{MLE}$ *

#7) $X_1,\dots X_n$ random sample from $N(0,\theta)$: $\theta\in(0,\infty)$ show that second order moment $\sum x_i^2$ is sufficient estimator of θ (a) joint pdf: $f(x_1,\dots x_n;\theta)=\prod_{i=1}^n\frac{1}{\sqrt{2\pi}\sqrt\theta}e^{-\frac{x_i^2}{2\theta}}\Rightarrow(2\pi\theta)^{-\frac{n}{2}}e^{\frac{-\sum x_i^2}{2\theta}}=\left[\theta^{-\frac{n}{2}}e^{-\frac{\sum x_i^2}{2\theta}}\right]\cdot(2\pi)^{-\frac{n}{2}}$
∴ By factorization Thm $\sum x_i^2$ is sufficient. * $g(\sum x_i^2;\theta)\cdot h(x_1,\dots x_n)$

#8) power funct. a test w/ H_0 about θ is $\pi(\theta)=P_\theta\{\text{reject }H_0\}$. let $\theta_0<\theta_1$: $\theta_1,\theta_2\in\mathbb{R}$
assume null is $H_0:\theta=\theta_0$. Sps tests 1,2,3 w/ π_1,π_2,π_3. $\pi_1(\theta_0)=\pi_2(\theta_0)=\pi_3(\theta_0)=\alpha\in(0,1)$
(a) which of 3 tests is size α? all 3, they all pass through (θ_0,α) *
(b) want to test $H_0:\theta=\theta_0$ vs. $H_1:\theta=\theta_1$ which is best most powerful? $\pi_1(\theta)$ (c) ditto $H_1:\theta\neq\theta_0$ $\pi_1(\theta)$ *



Test #2 prep | #2) $X_1,\dots X_n\sim f_\theta$: $f(x)=e^{-(x-\delta)}$: $x>\delta$ (a) consistent est. def.
(b) show smallest sample value is consistent est. of δ (b) $g(y_1)=nf(y_1)\left[\int_{y_1}^\infty f(x)dx\right]^{n-1}$: $y_1>\delta$
$=ne^{-n(y_1-\delta)}\Rightarrow$ use Y_1 estimate δ. Check $\forall\varepsilon>0$: $P(|y_1-\delta|\leq\varepsilon)=P(\delta-\varepsilon\leq y_1\leq\delta+\varepsilon)=P(\delta\leq y_1\leq\delta+\varepsilon)$
$=\int_\delta^{\delta+\varepsilon}ne^{-n(y_1-\delta)}dy_1$ set $t=y_1-\delta$ $=\int_0^\varepsilon ne^{-nt}dt=-e^{-nt}\Big|_0^\varepsilon=1-e^{-nt}\to1$ as $n\to\infty$ *

#4) 120 cheerleaders, 54 voice damage (a) $\forall\alpha\in(0,1)$ assert about maximum error obtaining sample proportion $\hat\theta=\frac{x}{n}$ as est.: $(1-\alpha)100\%$ (a) max. error $=z_{\frac{\alpha}{2}}\sqrt{\frac{\hat\theta(1-\hat\theta)}{n}}$ (b) 90% CI max error the $\frac{54}{120}=.45$ (b) $\alpha=0.1$, $z_\alpha=1.96$ * max $\varepsilon=1.96\sqrt{\frac{.45(1-.45)}{120}}$

#6) Sample from Poisson(λ) (b) Verify MVUE (c) CRLB
(a) definition of MVUE (a) $E(\hat\theta)=\theta$, $var(\hat\theta)$ = min. among unbiased est.
(c) mean of random sample (λ) is MVUE of λ (c) $X\sim Poisson(\lambda)$, $E(X)=\lambda$, $var(X)=\lambda$, $var(\bar X)=\frac{\lambda}{n}$, pdf $f(x;\lambda)=\frac{\lambda^x e^{-\lambda}}{x!}$
$\Rightarrow\log f(x;\lambda)=x\ln(\lambda)-\lambda-\ln(x!)$, $E\left[\frac{\partial\log f(x)}{\partial\lambda}\right]^2=E\left[\frac{x}{\lambda}-1\right]^2=E\left[\frac{x-\lambda}{\lambda}\right]^2=\lambda^{-2}E(X-\lambda)^2=\lambda^{-2}var(X)$
$=\lambda^{-2}\cdot\lambda=\lambda^{-1}\Rightarrow CRLB=\frac{1}{n\lambda^{-1}}=\frac{\lambda}{n}=var(\bar X)$ *

Test #1 prep | #1) X,Y,Z $\mu_X=2$, $\sigma_X^2=1$, $\mu_Y=-3$, $\sigma_Y^2=5$, $\mu_Z=1$, $\sigma_Z^2=3$ $cov(X,Y)=-2$, $cov(X,Z)=1$, $cov(Y,Z)=-1$
Find E( ) & var( ) of $W=3X-Y+Z$ (a) $E(W)=E(3X-Y+Z)=3\mu_X-\mu_Y+\mu_Z=3(2)-(-3)+1=10$ *
(b) $V(W)=V(3X)+V(-Y)+V(Z)+(2)(3)(-1)Cov(X,Y)+(2)(3)(1)Cov(X,Z)+(2)(-1)(1)cov(Y,Z)$
$=9\sigma_X^2+\sigma_Y^2+\sigma_Z^2+(-6)(-2)+(6)(1)+(-2)(-1)=9+5+3+12+6+2=37$ *

#2) $X_{11},X_{12},\dots X_{1n_1},X_{21},\dots X_{2n_2}$ independent r.v.s w/ $n_1,\mu_1,\sigma_1^2,n_2,\mu_2,\sigma_2^2$ (a) $E(\bar X_1-\bar X_2)=\mu_1-\mu_2\Rightarrow\bar X_i=\sum_{i=1}^{n_i}x_i$
$E(\bar X_1-\bar X_2)=E(\sum x_{i1})-E(\sum x_{i2})=\mu_1-\mu_2$ (b) $V(\hat\theta_1-\hat\theta_2)=\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}$ (b) $V(\hat\theta_1-\hat\theta_2)=V\left(\frac{1}{n_1}\sum x_{1i}\right)-V\left(\frac{1}{n_2}\sum x_{2i}\right)$
$\Rightarrow\frac{1}{n_1^2}V(\sum x_{1i})-\frac{1}{n_2^2}V(\sum x_{2i})=\frac{1}{n_1^2}n_1V(X_{1i})-\frac{1}{n_2^2}n_2V(X_{2i})=\frac{\theta_1(1-\theta_1)}{n_1}+\frac{\theta_2(1-\theta_2)}{n_2}=\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}$ *

#3) $X_1,\dots X_n\sim f_\theta$: $f(x)=e^{-(x-\delta)}$: $x>\delta$ (a) Find pdf of smallest sample value (a) $f(y_1)=nf(y_1)\left[\int_{y_1}^\infty f(x)dx\right]^{n-1}$
$=ne^{-(y_1-\delta)}\left[\int_{y_1}^\infty e^{-(x-\delta)}dx\right]^{n-1}\Rightarrow\int_{y_1}^\infty e^{-(x-\delta)}dx=-e^{-(x-\delta)}\Big|_{y_1}^\infty=e^{-(y_1-\delta)}\Rightarrow ne^{(y_1-\delta)}\left[e^{-(y_1-\delta)}\right]^{n-1}$

#4) Claim var $\sigma^2=4$ rejected if var of sample of 9 > 7.7535. Find prob. claim will be rejected (a) S^2, $n=9\Rightarrow\frac{(n-1)S^2}{\sigma^2}\sim\chi^2_{n-1}$
$P(S^2\geq7.7535)=P\left(\frac{(n-1)S^2}{\sigma^2}\geq\frac{8}{4}7.7535\right)=P(\chi^2\geq15.507)=0.05$

#5) Sample of $n=12$, $\bar x=27.8$, $s^2=3.24$ test claim on μ (a) choose statistic (a) $\frac{\bar X-\mu}{S/\sqrt n}\sim t_{n-1}$ (b) test claim that $\mu=25$
(b) $\frac{\bar X-\mu}{S/\sqrt n}=\frac{27.8-25}{1.8/\sqrt{12}}=5.3886>t_{11,0.05}=1.795$

#6) $X_1,\dots X_n$ random sample μ, $\sigma=3$ (a) use Chebyshev's determine smallest n must be taken to satisfy $P(|\bar X_n-\mu|\leq0.3)\geq0.95$ (b) use CLT to determine
(a) $P(|\bar X_n-\mu|\leq0.3)\geq1-\frac{Var(\bar X_n)}{0.3^2}\geq1-\frac{9}{n(0.09)}=1-\frac{100}{n}\geq0.95\Rightarrow\frac{100}{n}\leq0.05\Rightarrow n\geq\frac{100}{0.05}=2000$ (b) $P(|\bar Z|\leq\frac{0.3\sqrt n}{3})=0.95$
The smallest n: prob from part a holds (b) $P(|\bar X_n-\mu|\leq0.03)=P\left(\frac{|\bar X_n-\mu|}{\sigma/\sqrt n}\leq\frac{0.3}{\sigma/\sqrt n}\right)\approx$
$\Rightarrow\frac{\sqrt n}{10}=z_{0.95}=1.96\Rightarrow n=\lceil(19.6)^2\rceil=385$ *

Homework #6 | #1) Sample Poisson(λ) $H_0:\lambda_0=10$, $H_1:\lambda>10$ reject H_0 if $\bar X>12.5$, gets $\bar X=11.2$ (a) $\lambda=11.5$ (b) $\lambda=10$
(a) $\bar X$ estimates λ, $\lambda=11.5\not>12.5\Rightarrow H_0$ is false ∴ Type II error might occur. (b) $\lambda=10$ matches null ∴ no error will occur.

	$H_0=T$	$H_0=F$
reject	I	✓
fail to reject	✓	II

#2) $X\sim Geometric(\theta)$, $H_0:\theta=\theta_0$, $H_1:\theta=\theta_1$ $(\theta_1>\theta_0)$ null rejected iff $X\geq k$: $k\in\mathbb{Z}^+$ (a) Prob. of Type I (b) Prob II
(a) pmf: $P(X=k)=\theta(1-\theta)^{k-1}$ cdf: $P(X=k)=1-(1-\theta)^k$ $P(\text{Type I})=P(H_0\text{ rejected}|H_0\text{ is true})$
$\Rightarrow\alpha=P(X\geq k|\theta=\theta_0)=\sum\theta_0(1-\theta_0)^{k-1}$ Tail of geometric series starting at k ⇒ $\alpha=(1-\theta_0)^{k-1}$ *
$P(X\geq k|\theta=\theta_0)=1-P(X\leq k-1|\theta=\theta_0)=1-\left[1-(1-\theta_0)^{k-1}\right]=(1-\theta_0)^{k-1}$ (b) $P(\text{Type II})=P(H_0\text{ not rejected}|H_0\text{ is false})$
$\beta=P(X<k|\theta=\theta_1)=\sum_{i=1}^{k-1}(1-\theta_1)^{i-1}\theta_1$ sum of cdf upto $k-1$ ⇒ $\beta=1-(1-\theta_1)^{k-1}$ *

#3) $H_0:\beta=\beta_0$, $H_1:\beta=\beta_0+2$: Unif$[\alpha=0,\beta]$ ⇒ $f(x)=\frac{1}{\beta}$: $x\in[0,\beta]$ Reject H_0 iff $X>\beta_0+1$ (a) $P(\text{Type I})$ (b) $P(\text{Type II})$
(a) $\alpha=P(X>\beta_0+1|\beta=\beta_0)=0$ * since $\beta_0+1\notin[0,\beta_0]$ (b) $\beta=P(X\leq\beta_0+1|\beta=\beta_0+2)=\frac{\beta_0+1}{\beta_0+2}$ *

#4) random sample $\mu=0$, use Neyman-Pearson to construct most powerful critical region of size α test bb: $\sigma=\sigma_0$ to $H_1:\sigma=\sigma_1$, $\sigma_1>\sigma_0$
(a) $f(x;\sigma)=\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{x_i^2}{2\sigma^2}}$, $L(\sigma|x)=\prod f(x;\sigma)=\left(\frac{1}{\sqrt{2\pi}\sigma}\right)^n\exp\left[-\frac{1}{2\sigma^2}\sum x_i^2\right]\Rightarrow\frac{L_1}{L_0}=\left(\frac{\sigma_0}{\sigma_1}\right)^n\exp\left[\left(\frac{1}{2\sigma_0^2}-\frac{1}{2\sigma_1^2}\right)\sum_{i=1}^n x_i^2\right]$
inside exponent $-\frac{1}{2}\left(\frac{1}{\sigma_1^2}-\frac{1}{\sigma_0^2}\right)\sum x_i^2$ follow decreasing $T\sim\chi^2_n\Rightarrow P(T(x)<k|H_0)=\alpha\Rightarrow c=\left\{x\,|\,T(x)=\sum_{i=1}^n x_i^2<k\right\}$

Homework #5
#1) estimate w/ mean of 150 students est. mean score of a test. knowing $\sigma=9.4$ and prob 0.95 about maximum error?
(a) $n=150$, $\sigma=9.4$, $\alpha=0.05\Rightarrow1-\alpha=P\{|\bar X-\mu|<\max err\}=P\left(\left|\frac{\bar X-\mu}{\sigma/\sqrt n}\right|<\frac{\max err}{\sigma/\sqrt n}=z_{\frac{\alpha}{2}}\right)\Rightarrow z_{0.025}=1.96$
⇒ max error $=z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt n}=1.5043$ *

#2) Study of television viewing habits, mothers of teenagers. assume $\sigma=3.2$ how large is the sample w/ 95% CI that sample mean is 20 minutes? (a) $\alpha=0.05$, 20 min $=\frac{1}{3}$hr. $\Rightarrow1-\alpha=P(|\bar X-\mu|<\frac{1}{3})=P\left(\left|\frac{\bar X-\mu}{\sigma/\sqrt n}\right|<\frac{1}{3\sigma/\sqrt n}\right)$
$\Rightarrow z_{\frac{\alpha}{2}}=\frac{\sqrt n}{3\sigma}\Rightarrow n=\left(3\sigma z_{\frac{\alpha}{2}}\right)^2=\left[(3(3.2)(1.6)\right]^2=355$

#3) let $\bar X$ be the observed mean of sample n, from Normal(μ,σ) find n: $(\bar x-\frac{\sigma}{2},\bar x+\frac{\sigma}{2})$ approx 95% CI
(a) Known σ^2, $\alpha=0.05$, $z_{\frac{\alpha}{2}}=z_{0.025}=1.96\Rightarrow\mu\in\bar x\pm z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt n}$, $z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt n}=\frac{\sigma}{4}\Rightarrow n=(4z_{\frac{\alpha}{2}})^2=62$ *

#4) Two kinds of photocopying equipment 61 failures type1, avg of 80.7 min per repair std. dev 19.4 mins. Type 2 avg. 88, min/repair, std dev 18.8 min. find 99% CI for the difference between the averages
(a) (1) $n_1=61$, $\sigma_1=19.4$, $\bar X_1=80.2$ (2) $n_2=61$, $\sigma_2=18.8$, $\bar X_2=88.1$, $z_{0.05}=2.575\Rightarrow(\mu_1-\mu_2)\in(\bar X_1-\bar X_2)\pm z_{\frac{\alpha}{2}}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}$
$\Rightarrow-16.3\leq\mu_1-\mu_2\leq1.61$ *

#5) mine data $\begin{cases}\text{mine A: }8600, 8380, 8430, 7960, 8030\\ \text{mine B: }7710, 7890, 7920, 8270, 7860\end{cases}$ assuming independent from Normal pop. w/ equal variances find 99% CI of difference between the averages (a) pooled std dev. A: $n=5$, $\bar X_1=8260$, $S_1=251.8428$
B: $n_2=5$, $\bar X_2=7930$, $S_2=206.5188$ | df $=n_1+n_2-2$
$Sp=\sqrt{\frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}}=230.8259$, $\alpha=0.01$, $t_{\frac{\alpha}{2},8}=5.0411\Rightarrow\mu_1-\mu_2\in(\bar X_1-\bar X_2)\pm t_{\frac{\alpha}{2},df}S_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}$
$-404.327<\mu_1-\mu_2<1064.327$ *

#6) Among 100 fish, 18 inedible Construct 99% CI for true pop. (a) $n=100$, $\hat\theta=\frac{X}{n}=.18$, $\alpha=0.01$, $z_{0.05}=2.57$
$\Rightarrow\theta\in\hat\theta\pm z_{\frac{\alpha}{2}}\sqrt{\frac{\hat\theta(1-\hat\theta)}{n}}$ $0.0813\leq\theta\leq0.2787$ *

#7) length of skulls of 10 fossils w/ mean 5.68, std. dev. 0.29. assuming normal dist find 95% CI for (a) mean skull length (b) variance skull length
(a) $n=10$, $s=0.29$, $t^*=t_{\frac{\alpha}{2},n-1}=t_{0.025,9}=2.262\Rightarrow\mu\in\bar x\pm t^*\frac{s}{\sqrt n}\Rightarrow5.483\leq\mu\leq5.897$ * (b) $n=10$, $s=0.29$, $\alpha=0.05$
$\Rightarrow\chi^2_{\frac{\alpha}{2},n-1}=\chi^2_{0.025,9}=19.023$ & $\chi^2_{0.975,9}=2.70\Rightarrow\frac{(n-1)S^2}{\chi^2_{0.025}}<\sigma^2<\frac{(n-1)S^2}{\chi^2_{0.975,9}}\Rightarrow0.0398\leq\sigma^2\leq0.2803$ *

#8) Sample at Supermarket showed that 204 of 300 use coupons. Use a large sample CI for proportions w/ 95% of true prop.
(a) $n=300$, $\hat\theta=\frac{204}{300}$, $z_{\frac{\alpha}{2}}=z_{0.025}=1.96\Rightarrow\theta\in\hat\theta\pm z_{\frac{\alpha}{2}}\sqrt{\frac{\hat\theta(1-\hat\theta)}{n}}\Rightarrow\theta\in0.68\pm1.96\sqrt{\frac{(0.68)(0.32)}{300}}\Rightarrow\theta\in(0.627, 0.733)$ *

Homework #4 | #1) $X_1,\dots X_n$ random sample from Gamma($\alpha=3$), $\beta>0$ (a) method of moments (b) MLE (a) $\Gamma(n)=\int_0^\infty x^{n-1}e^{-ax}dx=\frac{\Gamma(n)}{a^n}$
$\Rightarrow f(x;\beta)=\frac{\beta^3}{\Gamma(3)}x^{3-1}e^{-\beta x}=\frac{\beta^3}{2}x^2e^{-\beta x}\Rightarrow E(X)=\frac{\beta^3}{2}\int_0^\infty x^3e^{-\beta x}dx=\frac{\beta^3}{2}\left(\frac{\Gamma(4)}{\beta^4}\right)=\frac{3}{\beta}\Rightarrow x=\frac{3}{\beta}\Rightarrow\beta=\frac{3}{\bar x}=\hat\beta_{mom}$
(b) $L(\beta)=\prod_{i=1}^n\frac{\beta^3}{2}x_i^2e^{-\beta x_i}=\frac{\beta^{3n}}{2n}\prod_{i=1}^n x_i^2e^{-\sum\beta x_i}\Rightarrow\ell(\beta)=\ln\left[\frac{\beta^{3n}}{2n}\prod_{i=1}^n x_i^2e^{-\sum\beta x_i}\right]=3n\ln(\beta)-n\ln(2)+\sum\ln(x_i^2)-\sum\beta x_i$
$\Rightarrow\frac{\partial\ell}{\partial\beta}=\frac{3n}{\beta}-\sum x_i=0\Rightarrow\beta=\frac{3n}{\sum x_i}=\frac{3n}{\bar x}$, $\hat\beta_{MLE}=\frac{3n}{\bar x}$ *

#2) random sample from Poisson(λ) use MLE for λ (a) $L=\prod_{i=1}^n f(x_i;\lambda)=\prod_{i=1}^n\frac{e^{-\lambda}\lambda^{x_i}}{x_i!}=\frac{e^{-\lambda}\lambda^{x_1}}{x_1!}\cdots\frac{e^{-\lambda}\lambda^{x_n}}{x_n!}$
$\Rightarrow\frac{e^{-n\lambda}}{x_1!\dots x_n!}\lambda^{\sum x_i}\Rightarrow\ell=\ln\left(\frac{e^{-n\lambda}}{x_1!\dots x_n!}\lambda^{\sum x_i}\right)=-n\lambda+\sum x_i\ln(\lambda)-\ln(x_1!\cdots x_n!)\Rightarrow$
$\Rightarrow\lambda=\frac{1}{n}\sum x_i=\bar x$ ∴ $\hat\lambda_{MLE}=\bar x$ *

#3) $X_1,\dots X_n$ constitute sample from Normal pop. $\mu=0$ show $\sum\frac{x_i^2}{n}$ is unbiased estimator of σ^2. (a) $E\left(\sum\frac{x_i^2}{n}\right)=\frac{1}{n}E\left(\sum x_i^2\right)$
$\Rightarrow\frac{1}{n}\sum E(x_i^2)=\frac{1}{n}$ ∵ $\mu=0$, $V(X_i)=\sigma^2$, $E(X_i^2)=V(X_i)+[E(X_i)]^2=\sigma^2+0\Rightarrow\frac{1}{n}\sum E(X_i^2)=\frac{1}{n}\sum\sigma^2=\frac{n}{n}\sigma^2=\sigma^2$
∴ $E\left(\sum\frac{x_i^2}{n}\right)=\sigma^2$ ∴ $\sum\frac{x_i^2}{n}$ is an unbiased estimator. *

#4) Show X r.v. of Binom(n,θ) show that $n\frac{X}{n}\left(1-\frac{X}{n}\right)$ is unbiased estimator of X. (a) $f(x)=\binom{n}{x}\theta^x(1-\theta)^{n-x}$, $E(X)=n\theta$, $V(X)=n\theta(1-\theta)$
$\Rightarrow E(X^2)=n\theta(1-\theta)+n\theta$ WTS: $E\left[n\frac{X}{n}\left(1-\frac{X}{n}\right)\right]\neq n\theta(1-\theta)\Rightarrow E\left[n\frac{X}{n}\left(1-\frac{X}{n}\right)\right]=\frac{1}{n}\left[nE\left(1-\frac{X}{n}\right)\right]=\frac{1}{n}E\left(n-\frac{nX}{n}\right)$
$\Rightarrow E\left[\frac{n}{n}-\frac{nX}{n^2}\right]=E\left[1-\frac{X}{n}\right]=1-E\left(\frac{X}{n}\right)=1-\frac{\theta}{n}\neq n\theta(1-\theta)$ *

#5) Show that the mean of sample n, is the minimum variance unbiased estimator of parameter λ (a) $X\sim Poisson(\lambda)$
$\Rightarrow E(X)=\lambda$, $\bar X$ is unbiased for any r.v. $X\sim Poisson$: $E(X)$ exists. ⇒ $\bar X=\lambda$ unbiased, $var(X)=\lambda\Rightarrow var(\bar X)=\frac{var(X)}{n}=\frac{\lambda}{n}$
$\Rightarrow\ln[f(x;\lambda)]=\ln\left[\frac{e^{-\lambda}\lambda^x}{x!}\right]=\ln(e^{-\lambda})+\ln(\lambda^x)-\ln(x!)=-\lambda+x\ln(\lambda)-\ln(x!)\Rightarrow\frac{\partial}{\partial\lambda}\ln f(x;\lambda)=-1+\frac{x}{\lambda}$
$\Rightarrow E\left[\left(\frac{\partial}{\partial\lambda}\ln f(x)\right)^2\right]=E\left[\left(-1+\frac{x}{\lambda}\right)^2\right]=E\left[\left(\frac{-\lambda+x}{\lambda}\right)^2\right]=\frac{1}{\lambda^2}E\left[(x-\lambda)^2\right]=\frac{1}{\lambda^2}Var(X)=\frac{1}{\lambda}\Rightarrow\frac{1}{nE\left[\left(\frac{\partial}{\partial\lambda}\ln f(x)\right)^2\right]}=\frac{1}{n\,1/\lambda}=\frac{\lambda}{n}$
∴ $var(\bar X)=\frac{1}{nE\left[\left(\frac{\partial}{\partial\lambda}\ln f(x)\right)^2\right]}$

#6) X_1,X_2,X_3 Sample $n=3$ from Normal w/ μ,σ^2 find efficiency of $T_1=\frac{X_1+2X_2+X_3}{4}$ relative to $T_2=\frac{X_1+X_2+X_3}{3}$
(a) $V(T_1)=V\left(\frac{X_1+2X_2+X_3}{4}\right)=\frac{1}{16}V(X_1+2X_2+X_3)=\frac{1}{16}\left[\sigma^2+4\sigma^2+\sigma^2\right]=\frac{6\sigma^2}{16}=\frac{3\sigma^2}{8}$, $V(T_2)=V\left(\frac{X_1+X_2+X_3}{3}\right)=\frac{3\sigma^2}{9}=\frac{\sigma^2}{3}$
T_1 relative to $T_2=V(T_2)/V(T_1)=\frac{\sigma^2/3}{3\sigma^2/8}=\frac{8}{9}$ *

#7) $X_1,\dots X_n\sim Exp(\theta)$, show $\bar X$ is a consistent est. of θ (a) $f(x)=\frac{1}{\theta}e^{-\frac{x}{\theta}}$, $E(X)=\theta$, $V(X)=\theta^2$ WTS: $\bar X=\frac{1}{n}\sum x_i$ consistent WTS: $V(\bar X)\to0$, $n\to\infty$
1) Unbiased ⇒ $E(\bar X)=\frac{1}{n}\sum E(X_i)=\frac{1}{n}\left[E(X_1)+\dots+E(X_n)\right]=\frac{1}{n}[\theta+\dots+\theta]=\frac{n\theta}{n}=\theta$, and $V(\bar X)=V\left(\frac{1}{n}\sum X_i\right)=\frac{1}{n^2}\left[V(X_1)+\dots V(X_n)\right]$
$\Rightarrow\frac{n\theta^2}{n^2}=\frac{\theta^2}{n}$, $\lim_{n\to\infty}V(\bar X)=\lim_{n\to\infty}\frac{\theta^2}{n}=0$ *

#8) $X_1,\dots X_n\sim$ Geometric show $y=X_1+X_2+\dots+X_n$ is a sufficient est of θ (a) $f(x)=\theta^x(1-\theta)^{x-1}$
$f(x_1,\dots x_n;\theta)=\prod_{i=1}^n f(x_i)=\prod_{i=1}^n\theta^{x_i}(1-\theta)^{x_i-1}=\frac{(1-\theta)^{\sum x_i}}{(1-\theta)^n}\cdot\theta^{\sum x_i}$ define $g:(u,\theta)\mapsto\frac{(1-\theta)^u}{(1-\theta)^n}\cdot\theta^u$ and $h(x_1,\dots x_n)=1$
denote $\hat\theta=\sum x_i\Rightarrow f(x_1,\dots x_n;\theta)=g(\hat\theta;\theta)h(x_1\dots x_n)$ by factorization thm. ⇒ $\hat\theta=\sum x_i$ is sufficient statistic *

Homework #3 | #2) $X_1,\dots X_n\sim$ geometric (a) method of moments (b) MLE (a) $f(x;\theta)=\theta(1-\theta)^x$, $E(X)=\sum x f(x)=\sum x\theta(1-\theta)^{x-1}$
$\Rightarrow\theta(1-\theta)\sum x(1-\theta)^{x-1}=\theta(1-\theta)\frac{\partial}{\partial\theta}\left[-\sum(1-\theta)^x\right]=\theta(1-\theta)\frac{\partial}{\partial\theta}\left[(1+(1-\theta)+(1-\theta)^2+\dots\right]=-\theta(1-\theta)\frac{\partial}{\partial\theta}\left[\frac{1}{1-(1-\theta)}\right]$ by geometric series
$=-\theta(1-\theta)\left(-\frac{1}{\theta^2}\right)=\frac{(1-\theta)}{\theta}=E(X)\Rightarrow\bar X=\frac{1-\theta}{\theta}\Rightarrow\bar X+1=\frac{1}{\theta}\Rightarrow\hat\theta=\frac{1}{\bar X+1}$ (b) $L=\prod_{i=1}^n\theta(1-\theta)^{x_i}=\theta^n\left[(1-\theta)^{\sum x_i}\right]$
$\Rightarrow\ell(\theta)=n\ln(\theta)+\sum x_i\ln(1-\theta)\Rightarrow\frac{n}{\theta}+\sum x_i\cdot\left(\frac{-1}{1-\theta}\right)\Rightarrow\frac{n}{\theta}-\frac{\sum x_i}{1-\theta}=0\Rightarrow\frac{n}{\theta}=\frac{\sum x_i}{1-\theta}\Rightarrow\frac{(1-\theta)n}{\theta}=\sum x_i\Rightarrow\frac{1-\theta}{\theta}=\frac{1}{n}\sum x_i$
∴ $\bar x=\frac{1-\theta}{\theta}$ or $\hat\theta=\frac{1}{1+\bar x}$ *

#3) $X_1,\dots X_n$ random sample from f_θ: $f(x)=\frac{2(\theta-x)}{\theta^2}$: $x\in(0,\theta)$ (a) $E(X)=\int_0^\theta x\frac{2(\theta-x)}{\theta^2}dx=\frac{1}{\theta^2}\left(\theta x^2-\frac{2x^3}{3}\right)\Big|_0^\theta=\frac{1}{3}\theta$ *
mom: $E(X)=\frac{1}{3}\theta=\bar X\Rightarrow\hat\theta_{mom}=3\bar X$

#4) Sample of size n from Normal with μ, find MLE of std. dev. σ (a) $L(\sigma^2)=\prod_{i=1}^n\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x_i-\mu)^2}{2\sigma^2}}=\frac{1}{(2\pi\sigma^2)^{n/2}}e^{-\frac{\sum(x_i-\mu)^2}{2\sigma^2}}$
$\Rightarrow\ell(\sigma^2)=-\frac{n}{2}\ln(2\pi)-\frac{n}{2}\ln(\sigma^2)-\frac{\sum(x_i-\mu)^2}{2\sigma^2}\Rightarrow\frac{\partial\ell}{\partial\sigma^2}=\frac{-n}{2\sigma^2}+\frac{\sum(x_i-\mu)^2}{2\sigma^4}=0\Rightarrow\sigma^2=\frac{\sum(x_i-\mu)^2}{n}\Rightarrow\sigma_{mle}=\sqrt{\frac{\sum(x_i-\mu)^2}{n}}$ *

#5) random sample from Beta pop. w/ $\beta=1$ use mom to estimate α (a) $X\sim Beta(\alpha,\beta)$: $\alpha,\beta>0$ $E(X)=\frac{\alpha}{\alpha+\beta}$, $\beta=1\Rightarrow E(X)=\frac{\alpha}{\alpha+1}=\bar X$
⇒ Solve for α: $\alpha=\bar X/(1-\bar X)$ *

Homework #2 | #1) experiment in Australia (both μ,σ^2 known) let $X_1\dots X_R$: R returns let $Z_i=\frac{(X_i-\mu)}{\sigma}\Rightarrow y=\frac{1}{k}\sum_{i=1}^R(X_i-\mu)^2=\frac{\sigma^2}{R}\sum z_i^2$
(a) name distribution of $W=Ry/\sigma^2$ (b) Sps $R=10$, $\sigma^2=0.09$ find c: $P(y\leq c)=0.99$ (a) $z_i=\frac{x_i-\mu}{\sigma}\sim N(0,1)$
$W=Ry/\sigma^2=\left(\frac{R}{\sigma^2}\right)\frac{\sigma^2}{R}\sum_{i=1}^R z_i^2$ where $\sum z_i^2\sim\chi^2_n\Rightarrow W$ has Chi-square dist (b) $R=10$, $\sigma^2=0.09$ * $\Rightarrow P(y\leq c)=P\left(\frac{R}{\sigma^2}y\leq\frac{R}{\sigma^2}c\right)$
$=P\left(W\leq\frac{10}{0.09}c\right)=0.99\Rightarrow\frac{10}{0.09}c=\chi^2_{10,0.01}=23.209\Rightarrow c=\frac{(0.09)\chi^2_{10,0.01}}{10}\approx0.209$ *

#2) from previous question. now sps $R=20$ from $X_1,\dots X_{20}$ let $\bar X=\frac{1}{20}\sum_{i=1}^{20}X_i$, find c: $P(\bar X_{20}\leq\mu+c\sigma)=0.95$ (a) $E(\bar X_{20})=\mu$, $V(\bar X_{20})=\frac{\sigma^2}{20}$
$\bar X\sim N(\mu,\frac{\sigma^2}{20})$, $P(\bar X_{20}\leq\mu+c\sigma)=P\left(\frac{\bar X_{20}-\mu}{\sigma/\sqrt{20}}\leq\frac{(\mu+c\sigma)-\mu}{\sigma/\sqrt{20}}\right)=P(Z\leq c\sqrt{20})=0.95$: $N(0,1)\Rightarrow\Phi(c\sqrt{20})=0.95$
$\Rightarrow c\sqrt{20}=z_{0.95}\approx1.65\Rightarrow c=\frac{z_{0.95}}{\sqrt{20}}\approx0.368$ *

Chebyshev's
#4) random sample $n=100$, infinite population with $\mu=75$, $\sigma^2=256$ (a) based on LLN w/ what prob. can we assert $\bar X\in(67,83)$ (b) on CLT
(a) $P(67<\bar X<83)=P(67-\mu<\bar X-\mu<83-\mu)=P(-8<\bar X-\mu<8)=P(|\bar X-\mu|<8)\geq1-\frac{\sigma^2}{n8^2}=1-\frac{256}{(100)(64)}=1-0.04=0.96$ *
(b) $P(67<\bar X<83)=P\left(\frac{67-\mu}{\sigma/\sqrt n}<z<\frac{83-\mu}{\sigma/\sqrt n}\right)=P(-5<Z<5)\overset{CLT}{\approx}P(-5<Z^*<5)\approx2.0499$ *

#5) sample of $n=225$ from Unif$[\alpha=24,\beta=48]$. Based on CLT find prob. that mean will be less than 35, $E(X)=\frac{\alpha+\beta}{2}$, $V(X)=\frac{(\beta-\alpha)^2}{12}$
(a) $X\sim Unif[\alpha,\beta]\Rightarrow E(X)=\frac{\alpha+\beta}{2}=\frac{24+48}{2}=36\Rightarrow P(\bar X<35)=P\left(\frac{\bar X-\mu}{\sigma/\sqrt n}<\frac{35-\mu}{\sigma/\sqrt n}\right)=P\left(Z<\frac{35-36}{\sqrt{48}/\sqrt{256}}\right)\approx P(Z<-2.17)$
$\overset{CLT}{\approx}P(Z^*<-2.17)\approx0.015$ *

Definitions

def Unbiased: A statistic $\hat\Theta$ is unbiased estimator of the parameter θ of a given dist iff $E(\hat\Theta)=\theta$

def: Asymptotically unbiased: letting $b_n(\Theta)=E(\hat\Theta)-\theta$ express the bias of an estimator θ based on a random sample of size n, from a given population, we say Θ is an asymptotically unbiased estimator iff $\lim_{n\to\infty}b_n(\Theta)=0$

def: Cramer-Rao: If $\hat\theta$ is an unbiased estimator of θ the variance of $\hat\theta$ must satisfy the inequality.
$var(\hat\theta)\geq\frac{1}{nE\left[\left(\frac{\partial\log f(x;\theta)}{\partial\theta}\right)^2\right]}$

hm: MVUE: If $\hat\theta$ is an unbiased estimator of θ ($E(\hat\theta)=\theta$) and $var(\hat\theta)=\frac{1}{nE\left[\left(\frac{\partial\log f(x;\theta)}{\partial\theta}\right)^2\right]}$ (CRLB)
Then $\hat\theta$ is minimum variance unbiased estimator of θ

def: Relative efficiency: two unbiased $\hat\theta_1,\hat\theta_2$ if $V(\hat\theta_1)<V(\hat\theta_2)\Rightarrow\hat\theta_1$ relatively more efficient than $\hat\theta_2$; $\frac{V(\hat\theta_1)}{V(\hat\theta_2)}$ $\hat\theta_2$ to $\hat\theta_1$

def: Consistent estimator: the statistic $\hat\Theta$ is a consistent estimator of parameter θ iff for each $c>0$ $\lim_{n\to\infty}P(|\hat\Theta-\theta|<c)=1$ ie $\hat\Theta\xrightarrow{P}\theta$ as $n\to\infty$

Thm: If $\hat\Theta$ is an unbiased estimator of parameter θ and $Var(\hat\Theta)\to0$ as $n\to\infty$ then $\hat\Theta$ is a consistent estimator of θ

def: Sufficient Estimator: The statistic $\hat\Theta=T(X)$ is a sufficient estimator of parameter θ of a given dist iff the conditional prob. dist./density of random samples X given T does not depend on the unknown parameter θ

Thm: factorization: Joint pdf $f(x_1,\dots x_n;\theta)=g(T(x);\theta)\cdot h(x_1,\dots x_n)$ or $\frac{f(x_1,\dots,x_n;\theta)}{g(T(x);\theta)}$ independent θ

Thm: Chebyshev's: r.v X will take a value $k>0$ std. dev. away from μ $P(|X-\mu|<k\sigma)\leq\frac{1}{k^2}$

Thm: LLN: $\forall\varepsilon>0$, prob. $\bar X$ will take a value within $[\mu-\varepsilon,\mu+\varepsilon]$ is at least $1-\frac{\sigma^2}{n\varepsilon^2}$ or $P(|\bar X-\mu|\leq\varepsilon)\geq1-\frac{\sigma^2}{n\varepsilon^2}$

Thm: CLT: $X_1,\dots X_n$ iid from same dist. w/ mean μ, var σ^2 then the limiting dist of $Z=\frac{\bar X-\mu}{\sigma/\sqrt n}$ is a std. normal as $n\to\infty$
$Z\xrightarrow{d}N(0,1)$ as $n\to\infty$

$Z=\frac{\bar X-\mu}{\sigma/\sqrt n}$	know σ^2,μ,n	$\Rightarrow Z\sim N(0,1)$ infer μ use $\bar X$
$T=\frac{\bar X-\mu}{S/\sqrt n}$	know S^2,μ,n	$\Rightarrow T\sim t_{n-1}$ infer μ, use $\bar X$
$\chi^2=\frac{(n-1)S^2}{\sigma^2}$	know S^2,σ^2	$\Rightarrow\chi^2\sim\chi^2_{n-1}$ infer σ^2 use S^2
$F=\frac{(S_1/\sigma_1^2)}{(S_2/\sigma_2^2)}$	know $S_1^2,S_2^2,\sigma_1^2,\sigma_2^2$	$\Rightarrow F\sim F_{S_1,S_2}$ infer ratio $\frac{\sigma_1^2}{\sigma_2^2}$

Practice #1 | #2) Independent Samples $n_1=30, n_2=50, \mu_1=78, \mu_2=75, \sigma_1^2=150, \sigma_2^2=200$ find prob. mean of first sample will exceed 2nd sample by 4.8 (a) $Y_1=\bar{X}_1-\bar{X}_2$, $\mu_y=\mu_1-\mu_2=78-75=3$, $\sigma_y^2=\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}=\frac{150}{30}+\frac{200}{50}=9$

$\Rightarrow P(\bar{X}_1-\bar{X}\geq 4.8)=P(Y\geq 4.8)=P\left(\frac{Y-\mu_y}{\sigma_y}\geq\frac{4.8-\mu_y}{\sigma_y}\right)=P\left(Z^*\geq\frac{4.8-3}{3}\right)=0.5-\Phi(0.6)=0.5-0.2257=0.2743$ *

#3) The claim that the variance of a Normal pop. is $\sigma^2=4$ is rejected if variance of random sample of $n=9$ exceeds 7.7535. Find prob. claim is to be rejected even though $\sigma^2=4$ (a) $P_{\sigma^2=4}(\text{Claim rejected})=P_{\sigma^2=4}(S^2>7.7535)=P\left(\frac{(9-1)S^2}{4}>\frac{(9-1)\,7.7535}{4}\right)=P(\chi_8^2>15.507)=0.05$ *

#4) random sample $n=25$ from Normal pop. has $\bar{x}=47, s=7$, based on sampling dist of sample mean w/ unknown variance can say given information supports conjecture that mean of pop. is $\mu=42$ (a) $|\bar{x}-42|=5$. $P_{\mu=42}(|\bar{X}-42|\geq 5)$

$=P\left(\left|\frac{\bar{X}-42}{S/\sqrt{25}}\right|\geq\frac{5}{s/\sqrt{25}}\right)=P(|t_{25-1}|\geq 3.571)=2\cdot P(t_{24}\geq 3.571)<2(0.005)=0.01$ * claim $\mu=42$ very small

#5) If S_1^2, S_2^2 are the variances of independent random samples of size $n_1=10, n_2=15$ from Normal pop. w/ equal variances (a) find $P(S_1^2/S_2^2<4.03)$ (a) $P_{\sigma_1^2=\sigma_2^2}\left(\frac{S_1^2}{S_2^2}<4.03\right)=P_{\sigma_1^2=\sigma_2^2}\left(\frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2}<\frac{4.03\sigma_2^2}{\sigma_1^2}\right)=P_{\sigma_1^2=\sigma_2^2}\left(\frac{\chi^2_{n_1-1}/(n_1-1)}{\chi^2_{n_2-1}/(n_2-1)}<4.03\right)$

$=P(F_{n_1-1,n_2-1}<4.03)=P(F_{9,14}<4.03)=1-P(F_{9,14}\geq 4.03)=1-0.1=0.99$ *

Practice #2 | #3) random sample $n=225$ from Exponential dist w/ $\theta=4$, w/ CLT what is prob. that mean will exceed 4.5

(a) let $X_1,\dots X_{225}\sim f_\theta: f(x)=\frac{1}{4}e^{-\frac{x}{4}}: x>0$, 0 otherwise, $E(X)=4, V(X)=4^2$, $P(\bar{X}>4.5)=P\left(\frac{\bar{X}-\mu_{\bar{X}}}{\sigma_{\bar{X}}}>\frac{4.5-4}{\sqrt{4^2/225}}\right)$

$\underset{\approx}{CLT}\ P(Z^*>1.875)=0.0304$ *

#5) Actual proportion of families in city, own their home is 0.70, 84 families $\sim$ Bernoulli$(n,\theta=0.70)$ w/ what prob. can we assert value of sample proportion $\hat{\Theta}\in(.64,.76)$ using (a) Chebyshev's (b) CLT (a) $X_1,\dots X_n$ (takes 1s, 0s) then sample proportion $\hat{\Theta}=\frac{1}{n}\sum X_i$ also $\mu_{\hat{\Theta}}=\theta=0.7$ and $\sigma_{\hat{\Theta}}^2=\frac{\sigma^2}{n}=\frac{\theta(1-\theta)}{84}=0.0025\Rightarrow P(.64<\hat{\Theta}<.76)$

$=P(0.64-0.7<\hat{\Theta}<0.76-0.7)=P(-0.06<\hat{\Theta}<0.06)=P(|\hat{\Theta}-\mu_{\hat{\Theta}}|<0.06)\geq 1-\frac{\sigma_{\hat{\Theta}}^2}{0.06^2}=1-\frac{0.0025}{0.06^2}=0.3056$

(b) $P(0.64<\hat{\Theta}<0.76)=P(|\hat{\Theta}-\mu_{\hat{\Theta}}|<0.06)=P\left\{\left|\frac{\hat{\Theta}-\mu_{\hat{\Theta}}}{\sigma_{\hat{\Theta}}}\right|<\frac{0.06}{\sigma_{\hat{\Theta}}}\right\}=P(|Z|<1.2)\underset{\approx}{CLT}P(|Z^*|<1.2)$

$=2\Phi(1.2)=2\cdot(3.849)=0.7698$ *

#6) Integrate the appropriate chi-square density to find the probability that the variance of a rand. sample of 5 from Normal pop. w/ $\sigma^2=25$, will fall between 20 and 30. (a) know $n=5, \sigma^2=25$, $P(20<S^2<30)=$

$=P\left\{\frac{(n-1)20}{\sigma^2}<\frac{(n-1)S^2}{\sigma^2}<\frac{(n-1)30}{\sigma^2}\right\}=P(3.2<\chi^2_{5-1}<4.8)=\frac{1}{2^{4/2}\Gamma(4/2)}\int_{3.2}^{4.8}x^{4/2-1}e^{-\frac{x}{2}}dx=\frac{1}{4(2-1)!}\int_{3.2}^{4.8}xe^{-\frac{x}{2}}dx$

$\Rightarrow u=x,\ v=-2e^{-\frac{x}{2}}$, $du=dx,\ dv=e^{-\frac{x}{2}}$ $\Rightarrow\frac{1}{4}\left(-2xe^{-\frac{x}{2}}\Big|_{3.2}^{4.8}+2\int_{3.2}^{4.8}e^{-\frac{x}{2}}dx\right)=-3.4e^{-2.4}+2.6e^{-1.6}\approx 0.216$ *

Practice #3 | #3) $X_1,\dots X_n\sim Exp(\theta)$ random sample show that $\bar{X}$ is a consistent estimator of θ (a) $E(X)=\theta, V(X)=\theta^2$

$\therefore E(\bar{X})=\theta=E(X)$ $\therefore$ unbiased estimator. also $V(\bar{X})=\frac{V(X)}{n}=\frac{\theta^2}{n}\to 0$ as $n\to\infty$ $\therefore\bar{X}$ is consistent *

#4) $X_1,\dots X_n\sim Exp(\theta)$ show that $\bar{X}$ is a sufficient estimator of θ (a) $X\sim Exp(\theta)$, $X_1,\dots X_n$ sample $\forall i\in\{1,\dots n\}\ X_i>0$

joint pdf: $f(x_1,\dots x_n;\theta)=\prod_{i=1}^n f_X(x_i)=\prod_{i=1}^n\frac{1}{\theta}e^{\frac{-x_i}{\theta}}=\frac{1}{\theta^n}e^{-\frac{1}{\theta}\sum x_i}=\frac{1}{\theta^n}e^{-\frac{n}{\theta}\bar{x}}\Rightarrow$ define $g:(u,\theta)\mapsto\frac{1}{\theta^n}e^{-\frac{n}{\theta}u}$

and $h(x_1,\dots x_n;\theta)\equiv 1\Rightarrow f(x_1,\dots x_n;\theta)=g(\bar{x},\theta)h(x_1,\dots x_n)$ By factorization thm $\bar{X}$ is sufficient statistic *

#5) $X_1,\dots X_n\sim Poisson(\lambda)$ sample obtain estimator of λ for (a) mom (b) MLE (a) $E(X)=\lambda$, let $\bar{X}=E(X)=\lambda\Rightarrow\hat{\lambda}_{mom}=\bar{X}$ *

(b) MLE $L(\lambda)=\prod_{i=1}^n f_X(x_i)=\prod_{i=1}^n e^{-\lambda}\frac{\lambda^{x_i}}{x_i!}=e^{-n\lambda}\frac{\lambda^{\sum x_i}}{\prod x_i!}\Rightarrow\ell(\lambda)=-n\lambda+\ln(\lambda)\sum x_i-\ln\left(\prod_{i=1}^n x_i!\right)$

$\Rightarrow\frac{\partial\ell}{\partial\lambda}=-n+\frac{\sum x_i}{\lambda}=0$ solve for $\lambda\Rightarrow\hat{\lambda}_{MLE}=\frac{1}{n}\sum x_i=\bar{x}$

#6) random sample from Normal known μ, and finite variance σ^2 show $\sum(X_i-\mu)^2$ is unbiased estimator

(a) $E\left[\frac{1}{n}\sum(X_i-\mu)^2\right]=\frac{1}{n}\sum E[(X_i-\mu)^2]=\frac{1}{n}\sum V(X)=\frac{1}{n}n\sigma^2=\sigma^2$ *

#7) $X_1,\dots X_n$ random sample from Normal pop w/ $\mu=0$, show $\sum\frac{X_i^2}{n}$ is unbiased estimator of σ^2 (a) $X\sim N(\mu=0,\sigma^2)$

$\Rightarrow E\left[\sum\frac{X_i^2}{n}\right]=\frac{1}{n}\sum E(X_i^2)=\frac{n}{n}E(X^2)\Rightarrow E(X^2)=E[(X-\mu)^2]=Var(X)=\sigma^2$ * $\because\mu=0$

#8) show that if $\hat{\theta}$ is an unbiased estimator of θ and $Var(\hat{\theta})\neq 0$ the $\hat{\theta}^2$ is not an unbiased est. of θ

(a) since $V(\hat{\theta})=E(\hat{\theta}^2)-[E(\hat{\theta})]^2$ assuming unbiasedness $E(\hat{\theta})=\theta$, and proper randomness $V(\hat{\theta})\neq 0\Rightarrow E(\hat{\theta}^2)-\theta^2\neq 0$

$\Leftrightarrow E(\hat{\theta}^2)\neq\theta^2\Rightarrow$ Not an unbiased estimator of θ. *

Practice #4 | #1) Among 500 marriage license app. 48 women at least one year older than men. Chosen 400 6 years later, 68. Construct 99% CI difference between true proportions. (a) $n_1=500, n_2=400$, $\hat{\theta}_1=48/500=0.096$, $\hat{\theta}_2=68/400=0.170$, $Z_{\frac{\alpha}{2}}=Z_{0.05}=2.575$

$\Rightarrow\theta_1-\theta_2\in(\hat{\theta}_1-\hat{\theta}_2)\pm z_{\frac{\alpha}{2}}\sqrt{\frac{\hat{\theta}_1(1-\hat{\theta}_1)}{n_1}+\frac{\hat{\theta}_2(1-\hat{\theta}_2)}{n_2}}\Rightarrow\theta_1-\theta_2\in(0.096-0.17)\pm(2.575)\sqrt{\frac{0.096(0.904)}{500}+\frac{0.17(0.830)}{400}}$

$\Rightarrow\theta_1-\theta_2\in(-.133,-.015)$ *

#2) w/ reference to previous question 4 can we say w/ 98% conf. about max error about difference among sample prop. and true?

(a) $\alpha=0.02$, $\Rightarrow Z_{\frac{\alpha}{2}}=Z_{0.01}=$ $1-\alpha=P\{|(\theta_1-\theta_2)-(\hat{\theta}_1-\hat{\theta}_2)|\}\approx P\left\{|Z|<\frac{\text{max. error}}{\sqrt{\frac{\hat{\theta}_1(1-\hat{\theta}_1)}{n_1}+\frac{\hat{\theta}_2(1-\hat{\theta}_2)}{n_2}}}\right\}$

$\Rightarrow Z_{\frac{\alpha}{2}}=\frac{\text{max. error}}{\sqrt{\frac{\hat{\theta}_1(1-\hat{\theta}_1)}{n_1}+\frac{\hat{\theta}_2(1-\hat{\theta}_2)}{n_2}}}\Rightarrow$ * $\text{max. error}=Z_{\frac{\alpha}{2}}\sqrt{\frac{\hat{\theta}_1(1-\hat{\theta}_1)}{n_1}+\frac{\hat{\theta}_2(1-\hat{\theta}_2)}{n_2}}$

Practice #5 | #1) A single observation of a r.v. having hypergeometric dist w/ $N=7, n=2$, used to test null hyp. of $k=2$ vs alternative $k=4$. If null hypothesis is rejected iff value of the r.v. is 2. find prob. of type I, and type II errors.

(a) population of $N=7$ objects and choose 2 of them. Among 7, have "good" and remaining $N-k=7-k$ "not good". If X is the # of "good" objects among $n=2$, $\Rightarrow P\{X=x\}=\frac{\binom{k}{x}\binom{N-k}{n-x}}{\binom{N}{n}}=\frac{\binom{k}{x}\binom{7-k}{2-x}}{\binom{7}{2}}$, $x=0,1,2$, since rejection region is based on observed value of X,

$\Rightarrow\alpha=\text{Type I error}=P_{k=2}\{X=2\}=\frac{\binom{2}{2}\cdot 1}{\binom{7}{2}}=\frac{1}{\frac{7\cdot 6}{2}}=\frac{1}{21}$

$\Rightarrow\beta=\text{Type II error}=P_{k\neq 4}\{X\neq 2\}=1-P_{k=2}\{X=2\}=1-\frac{\binom{4}{2}\cdot 1}{\binom{7}{2}}=\frac{5}{7}$

#2) X_1, X_2 constitute a random sample of size 2, from $f_\theta: f(x)=\theta x^{\theta-1}: x\in(0,1)$ If critical region $X_1X_2\geq\frac{3}{4}$, used to test null hypothesis $\theta=1$, vs alternative $\theta=2$ what's the power of test at $\theta=2$ (a) critical region $C=\{(x_1,x_2)\,|\,x_1x_2\geq\frac{3}{4}\}\Rightarrow$ power $\theta=2$

$=\pi(\theta)\big|_{\theta=2}=P_{\theta=2}\{(X_1,X_2)\in C\}=P_{\theta=2}\{X_1X_2\geq\frac{3}{4}\}$ joint pdf assuming independence $f(x_1,x_2)=f_X(x_1)f_X(x_2)=\theta^2(x_1x_2)^{\theta-1}$

for $0<x_1$ and $x_2<1$ $f(x_1,x_2)=0\Rightarrow\pi(2)=P_{\theta=2}\{X_1X_2\geq\frac{3}{4}\}=\int_{\frac{3}{4}}^1\int_{\frac{3}{4x_1}}^1 2^2(x_1x_2)^{2-1}dx_1dx_2=4\int_{\frac{3}{4}}^1 x_1\int_{\frac{3}{4x_1}}^1 x_2dx_2dx_1$

$\Rightarrow 4\int_{\frac{3}{4}}^1 x_1\frac{x_2^2}{2}\Big|_{x_2=\frac{3}{4x_1}}^{x_2=1}dx_1=\frac{4}{2}\int_{\frac{3}{4}}^1 x_1\left(1^2-\left(\frac{3}{4x_1}\right)^2\right)dx_1=2\int_{\frac{3}{4}}^1 x_1\left(1-\frac{9}{16x_1^2}\right)dx_1=2\int_{\frac{3}{4}}^1 x_1-\frac{9}{16x_1}dx_1=2\left(\frac{x_1^2}{2}-\frac{9}{16}\ln|x_1|\right)\Big|_{\frac{3}{4}}^1$

$=x_1^2\Big|_{\frac{3}{4}}^1-\frac{9}{8}\ln|x_1|\Big|_{\frac{3}{4}}^1=\left(1-\frac{9}{16}\right)-\frac{9}{8}\left(0-\ln\left(\frac{3}{4}\right)\right)\approx 0.1139$ *

#3) random sample of size n, used to test null hypothesis of θ, of an Exponential(θ) equals θ_0 vs alternative doesn't equal θ_0 (a) Find an expression for likelihood ratio statistic (b) use result of part (a) to show critical region of likelihood ratio statistic can be written as $\bar{x}\cdot\exp\left\{\frac{-\bar{x}}{\theta_0}\right\}\leq k$ (a) likelihood $L(\theta)=\prod_{i=1}^n f(x_i)=\prod_{i=1}^n\frac{1}{\theta}e^{\left(-\frac{x_i}{\theta}\right)}=\frac{1}{\theta^n}e^{-\frac{1}{\theta}\sum x_i}=\frac{1}{\theta^n}e^{\frac{-n\bar{x}}{\theta}}$

$\Rightarrow$ likelihood ratio statistic $\Lambda(X_1,\dots X_n)=\frac{\max_{\theta\in\{\theta_0\}}L(\theta)}{\max_{\theta>0}L(\theta)}=\frac{L(\theta_0)}{L(\hat{\theta}_{MLE})}$ want to maximize $L(\theta)$ at $\theta_0\in(0,\infty)$

$\Rightarrow\ln L(\theta)=-n\ln(\theta)-\frac{n\bar{x}}{\theta}$ easier to find $\ln L(\theta)$ since increasing funct.

$\Rightarrow\frac{\partial\ell}{\partial\theta}=\frac{-n}{\theta}+\frac{n\bar{x}}{\theta^2}\overset{set}{=}0\Rightarrow$ mult. $\frac{\theta^2}{n}\Rightarrow-\theta+\bar{x}=0\Rightarrow\hat{\theta}_{MLE}=\bar{x}\Rightarrow\Lambda(x_1,\dots x_n)=\frac{\frac{1}{\theta_0^n}e^{-\frac{n\bar{x}}{\theta_0}}}{\frac{1}{\bar{x}^n}e^{-\frac{n\bar{x}}{\bar{x}}}}=\frac{\bar{x}^n}{\theta_0^n}e^{-n\left(\frac{\bar{x}}{\theta_0}-1\right)}$

$=\frac{\bar{x}^n}{\theta_0^n}e^{\frac{n}{\theta_0}(\bar{x}-\theta_0)}$ (b) Critical region $\{\Lambda\leq k\}$ for some k. $\Rightarrow\frac{\bar{x}^n}{\theta_0^n}\left(e^{\frac{-1}{\theta_0}(\bar{x}-\theta_0)}\right)^n\leq k\Rightarrow\frac{\bar{x}}{\theta_0}e^{-\frac{1}{\theta_0}(\bar{x}-\theta_0)}\leq\sqrt[n]{k}$

$\Rightarrow\frac{\bar{x}}{\theta_0}e^{\frac{-\bar{x}}{\theta_0}}e\leq\sqrt[n]{k}\Rightarrow\bar{x}e^{-\frac{\bar{x}}{\theta_0}}\leq\frac{\theta_0\sqrt[n]{k}}{e}\triangleq K$ *

#4) A city police considering replacement of tires. If μ_1 is avg # of miles that old last, μ_2 avg. # of new tires will last, null hyp $\mu_1=\mu_2$

(a) what alternative hypothesis should department use if does not want to use new tires unless definitely proved to get better mileage?

(b) " if it is anxious to get new tires unless they actually give less milage than old? old tires kept if null hyp. rejected

(c) " so that rejection of null hyp. can lead to either keeping old or buying new tires?

(a) alternative: $H_1:\mu_2>\mu_1$ (b) $H_1:\mu_2<\mu_1$ (c) $H_1:\mu_1\neq\mu_2$

Practice #6 | #1) Single observation of r.v. w/ Geometric dist used to test null hyp. $\theta=\theta_0$ vs. alternative $\theta=\theta_1>\theta_0$. If null hyp. is rejected $\Leftrightarrow$ observed value of r.v. $\geq k: k\in\mathbb{Z}^+$

(a) find expressions for prob. of type I and II (a) $\alpha=(1-\theta_0)^{k-1}$, $\beta=1-(1-\theta_1)^{k-1}$ *

#2) $X_1, X_2\sim N(\mu,\sigma^2=1)$ If null $\mu=\mu_0$ rejected in favor of alternative $\mu=\mu_1>\mu_0$: $\bar{X}>\mu_0=1$

(a) what's the size of critical region? (a) $\alpha=0.08$ *

#3) random sample size n, $\sim$Exponential test null $\theta=\theta_0$ vs alternative $\theta=\theta_1>\theta_0$. use Neyman-Pearson to find most powerful region of size α. (a) $\sum X_i\geq K$: K determined by size α, and use $\sum X_i\sim$ Gamma$(\alpha=n,\beta=\theta_0)$

#4) from Normal pop. $\mu=0$ use Neyman-Pearson construct most powerful region α, test $\sigma=\sigma_0$ vs. $\sigma=\sigma_1>\sigma_0$

(a) $\sum X_i\geq K$ determined by α, use $\frac{1}{n}\sum X_i^2\sim\chi_n^2$ dist *

#5) Single obs. used to test null hyp. of mean waiting time between tremor at seismological station (mean of exponential pop) is $\theta=10$ hours. against alternative $\theta\neq 10$. If null is rejected iff the obs. values is less than 8 or greater than 12. find (a) Prob. of Type I (b) Prob. of type II errors when $\theta=2,4,6,8,12,16$ and 20.

(a) 0.852 (b) 0.016, 0.089, 0.129, 0.145, 0.144, 0.134, 0.122 *