Final EXAM- ADDENDUM

Method of moments Estimation:

Equate the population and sample moments to estimate the parameters

Moment (population): $\mu_i = E(X^i)$ Central Moment (population): $\mu'_i = E(X - \mu_1)^i$

Moment (sample): $m_i = (\sum X^i) / n$ Central Moment (sample): $m'_i = \sum(X - m_1)^i / n$

Method of Max Likelihood:

Find the value for parameter that maximizes log-likelihood by equating its derivative to 0

$$\frac{d}{d\theta}(\ln L(X_1, X_2, \dots X_n)) = \frac{d}{d\theta}(\ln(P(X_1)P(X_2)\dots P(X_n))) = 0$$

Confidence Intervals:

Confidence interval, Normal distribution	If parameter θ has an unbiased, Normally distributed estimator $\hat{\theta}$, then $\hat{\theta} \pm z_{\alpha/2} \cdot \sigma(\hat{\theta}) = \left[\hat{\theta} - z_{\alpha/2} \cdot \sigma(\hat{\theta}),\ \hat{\theta} + z_{\alpha/2} \cdot \sigma(\hat{\theta})\right]$ is a $(1-\alpha)100\%$ confidence interval for θ. If the distribution of $\hat{\theta}$ is *approximately* Normal, we get an *approximately* $(1-\alpha)100\%$ confidence interval.

If we do not know the population std. dev. but we know the n is large, then σ(θ) can be replaced by s(θ)

Confidence interval for the difference of means; known standard deviations	$\bar{X} - \bar{Y} \pm z_{\alpha/2}\sqrt{\frac{\sigma_X^2}{n} + \frac{\sigma_Y^2}{m}}$
Confidence interval for a population proportion	$\hat{p} \pm z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$
Confidence interval for the difference of proportions	$\hat{p}_1 - \hat{p}_2 \pm z_{\alpha/2}\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$

Common Z values:

$$z_{0.10} = 1.282, \quad z_{0.05} = 1.645, \quad z_{0.025} = 1.960$$
$$z_{0.01} = 2.326, \quad z_{0.005} = 2.576.$$

Can also obtained from Z-table or from T-table with v = ∞

If n is small,

Confidence interval for the mean; σ is unknown	$\bar{X} \pm t_{\alpha/2} \frac{s}{\sqrt{n}}$ where $t_{\alpha/2}$ is a critical value from T-distribution with $n-1$ degrees of freedom
Confidence interval for the difference of means; equal, unknown standard deviations	$\bar{X} - \bar{Y} \pm t_{\alpha/2}\, s_p \sqrt{\frac{1}{n} + \frac{1}{m}}$ where s_p is the *pooled standard deviation*, a root of the pooled variance in (9.11) and $t_{\alpha/2}$ is a critical value from T-distribution with $(n+m-2)$ degrees of freedom

Pooled Std. Deviation:

$$s_p^2 = \frac{\sum_{i=1}^{n}(X_i - \bar{X})^2 + \sum_{i=1}^{m}(Y_i - \bar{Y})^2}{n+m-2} = \frac{(n-1)s_X^2 + (m-1)s_Y^2}{n+m-2}. \qquad (9.11)$$

Confidence interval for the difference of means; unequal, unknown standard deviations	$\bar{X} - \bar{Y} \pm t_{\alpha/2} \sqrt{\frac{s_X^2}{n} + \frac{s_Y^2}{m}}$ where $t_{\alpha/2}$ is a critical value from T-distribution with ν degrees of freedom given by formula (9.12)

Statterthwaite approximation of degrees of freedom (9.12) [Round up to nearest integer]:

$$\nu = \frac{\left(\frac{s_X^2}{n} + \frac{s_Y^2}{m}\right)^2}{\frac{s_X^4}{n^2(n-1)} + \frac{s_Y^4}{m^2(m-1)}}$$

Hypothesis Testing (Z-tests):

Right Tailed Test (H_A: $\theta > \theta_0$): $\begin{cases} \text{reject } H_0 & \text{if } \quad Z \geq z_\alpha \\ \text{accept } H_0 & \text{if } \quad Z < z_\alpha \end{cases}$

Left Tailed Test (H_A: $\theta < \theta_0$): $\begin{cases} \text{reject } H_0 & \text{if } \quad Z \leq -z_\alpha \\ \text{accept } H_0 & \text{if } \quad Z > -z_\alpha \end{cases}$

Two Tailed Test (H_A: $\theta \neq \theta_0$): $\begin{cases} \text{reject } H_0 & \text{if } \quad |Z| \geq z_{\alpha/2} \\ \text{accept } H_0 & \text{if } \quad |Z| < z_{\alpha/2} \end{cases}$

Hypothesis Testing (t-tests):

For a **right-tail alternative**,

$$\begin{cases} \text{reject } H_0 & \text{if} \quad t \ge t_\alpha \\ \text{accept } H_0 & \text{if} \quad t < t_\alpha \end{cases}$$

For a **left-tail alternative**,

$$\begin{cases} \text{reject } H_0 & \text{if} \quad t \le -t_\alpha \\ \text{accept } H_0 & \text{if} \quad t > -t_\alpha \end{cases}$$

For a **two-sided alternative**,

$$\begin{cases} \text{reject } H_0 & \text{if} \quad |t| \ge t_{\alpha/2} \\ \text{accept } H_0 & \text{if} \quad |t| < t_{\alpha/2} \end{cases}$$

Summary of Z - Tests

Null hypothesis	Parameter, estimator	If H_0 is true:		Test statistic
H_0	θ, $\hat{\theta}$	$\mathbf{E}(\hat{\theta})$	$\mathrm{Var}(\hat{\theta})$	$Z = \dfrac{\hat{\theta} - \theta_0}{\sqrt{\mathrm{Var}(\hat{\theta})}}$
One-sample Z-tests for means and proportions, based on a sample of size n				
$\mu = \mu_0$	μ, $\bar{X}$	μ_0	$\dfrac{\sigma^2}{n}$	$\dfrac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}$
$p = p_0$	p, $\hat{p}$	p_0	$\dfrac{p_0(1-p_0)}{n}$	$\dfrac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$
Two-sample Z-tests comparing means and proportions of two populations, based on independent samples of size n and m				
$\mu_X - \mu_Y = D$	$\mu_X - \mu_Y$, $\bar{X} - \bar{Y}$	D	$\dfrac{\sigma_X^2}{n} + \dfrac{\sigma_Y^2}{m}$	$\dfrac{\bar{X} - \bar{Y} - D}{\sqrt{\frac{\sigma_X^2}{n} + \frac{\sigma_Y^2}{m}}}$
$p_1 - p_2 = D$	$p_1 - p_2$, $\hat{p}_1 - \hat{p}_2$	D	$\dfrac{p_1(1-p_1)}{n} + \dfrac{p_2(1-p_2)}{m}$	$\dfrac{\hat{p}_1 - \hat{p}_2 - D}{\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n} + \frac{\hat{p}_2(1-\hat{p}_2)}{m}}}$
$p_1 = p_2$	$p_1 - p_2$, $\hat{p}_1 - \hat{p}_2$	0	$p(1-p)\left(\dfrac{1}{n} + \dfrac{1}{m}\right)$, where $p = p_1 = p_2$	$\dfrac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n} + \frac{1}{m}\right)}}$ where $\hat{p} = \dfrac{n\hat{p}_1 + m\hat{p}_2}{n+m}$

Summary of t-tests

Hypothesis H_0	Conditions	Test statistic t	Degrees of freedom
$\mu = \mu_0$	Sample size n; unknown σ	$t = \dfrac{\bar{X} - \mu_0}{s/\sqrt{n}}$	$n - 1$
$\mu_X - \mu_Y = D$	Sample sizes n, m; unknown but equal standard deviations, $\sigma_X = \sigma_Y$	$t = \dfrac{\bar{X} - \bar{Y} - D}{s_p\sqrt{\frac{1}{n} + \frac{1}{m}}}$	$n + m - 2$
$\mu_X - \mu_Y = D$	Sample sizes n, m; unknown, unequal standard deviations, $\sigma_X \neq \sigma_Y$	$t = \dfrac{\bar{X} - \bar{Y} - D}{\sqrt{\frac{s_X^2}{n} + \frac{s_Y^2}{m}}}$	Satterthwaite approximation, formula (9.12)

P Values (Reject H_0 if $P < 0.01$, Accept H_0 if $P > 0.1$, Not enough evidence otherwise):

P values for Z tests:

Hypothesis H_0	Alternative H_A	P-value	Computation
$\theta = \theta_0$	right-tail $\theta > \theta_0$	$\boldsymbol{P}\{Z \geq Z_{\text{obs}}\}$	$1 - \Phi(Z_{\text{obs}})$
	left-tail $\theta < \theta_0$	$\boldsymbol{P}\{Z \leq Z_{\text{obs}}\}$	$\Phi(Z_{\text{obs}})$
	two-sided $\theta \neq \theta_0$	$\boldsymbol{P}\{\lvert Z\rvert \geq \lvert Z_{\text{obs}}\rvert\}$	$2(1 - \Phi(\lvert Z_{\text{obs}}\rvert))$

P values for t tests:

Hypothesis H_0	Alternative H_A	P-value	Computation
$\theta = \theta_0$	right-tail $\theta > \theta_0$	$\boldsymbol{P}\{t \geq t_{\text{obs}}\}$	$1 - F_\nu(t_{\text{obs}})$
	left-tail $\theta < \theta_0$	$\boldsymbol{P}\{t \leq t_{\text{obs}}\}$	$F_\nu(t_{\text{obs}})$
	two-sided $\theta \neq \theta_0$	$\boldsymbol{P}\{\lvert t\rvert \geq \lvert t_{\text{obs}}\rvert\}$	$2(1 - F_\nu(\lvert t_{\text{obs}}\rvert))$

Confidence Intervals (variance):

Confidence interval for the variance	$$\left[\frac{(n-1)s^2}{\chi^2_{\alpha/2}},\ \frac{(n-1)s^2}{\chi^2_{1-\alpha/2}}\right]$$
Confidence interval for the standard deviation	$$\left[\sqrt{\frac{(n-1)s^2}{\chi^2_{\alpha/2}}},\ \sqrt{\frac{(n-1)s^2}{\chi^2_{1-\alpha/2}}}\right]$$

Hypothesis tests for variance (can also be used for Std Dev by conv question to variance):

Null Hypothesis	Alternative Hypothesis	Test statistic	Rejection region	P-value
$\sigma^2 = \sigma_0^2$	$\sigma^2 > \sigma_0^2$	$\frac{(n-1)s^2}{\sigma_0^2}$	$\chi^2_{\text{obs}} \geq \chi^2_\alpha$	$\boldsymbol{P}\left\{\chi^2 \geq \chi^2_{\text{obs}}\right\}$
	$\sigma^2 < \sigma_0^2$		$\chi^2_{\text{obs}} \leq \chi^2_{1-\alpha}$	$\boldsymbol{P}\left\{\chi^2 \leq \chi^2_{\text{obs}}\right\}$
	$\sigma^2 \neq \sigma_0^2$		$\chi^2_{\text{obs}} \geq \chi^2_{\alpha/2}$ or $\chi^2_{\text{obs}} \leq \chi^2_{1-\alpha/2}$	$2\min\left(\boldsymbol{P}\left\{\chi^2 \geq \chi^2_{\text{obs}}\right\}, \boldsymbol{P}\left\{\chi^2 \leq \chi^2_{\text{obs}}\right\}\right)$

Testing ratio of Variances (can also be used for Std Dev by conv question to variance):

Null Hypothesis $H_0: \frac{\sigma_X^2}{\sigma_Y^2} = \theta_0$		Test statistic $F_{\text{obs}} = \frac{s_X^2}{s_Y^2}/\theta_0$
Alternative Hypothesis	Rejection region	P-value Use F$(n-1, m-1)$ distribution
$\frac{\sigma_X^2}{\sigma_Y^2} > \theta_0$	$F_{\text{obs}} \geq F_\alpha(n-1, m-1)$	$\boldsymbol{P}\left\{F \geq F_{\text{obs}}\right\}$
$\frac{\sigma_X^2}{\sigma_Y^2} < \theta_0$	$F_{\text{obs}} \leq 1/F_\alpha(m-1, n-1)$	$\boldsymbol{P}\left\{F \leq F_{\text{obs}}\right\}$
$\frac{\sigma_X^2}{\sigma_Y^2} \neq \theta_0$	$F_{\text{obs}} \geq F_{\alpha/2}(n-1, m-1)$ or $F_{\text{obs}} \leq 1/F_{\alpha/2}(m-1, n-1)$	$2\min\left(\boldsymbol{P}\left\{F \geq F_{\text{obs}}\right\}, \boldsymbol{P}\left\{F \leq F_{\text{obs}}\right\}\right)$

Bayesian Statistics

Given a prior distribution π(θ) and a model for some observations $f(x|\theta) = f(x_1, x_2, x_3, \dots x_n|\theta)$ the posterior distributions $\pi(\theta|x)$ is given by

Posterior distribution

$$\boxed{\pi(\theta|\boldsymbol{x}) = \pi(\theta|\boldsymbol{X} = \boldsymbol{x}) = \frac{f(\boldsymbol{x}|\theta)\pi(\theta)}{m(\boldsymbol{x})}.}$$

Where

Marginal distribution of data

$$m(\boldsymbol{x}) = \sum_{\theta} f(x|\theta)\pi(\theta)$$

for discrete prior distributions π

$$m(\boldsymbol{x}) = \int_{\theta} f(x|\theta)\pi(\theta)d\theta$$

for continuous prior distributions π

This holds true for pmf and for pdfs.

Conjugate families for Bayesian statistics

Model $f(\boldsymbol{x}\mid\theta)$	Prior $\pi(\theta)$	Posterior $\pi(\theta\mid\boldsymbol{x})$
Poisson(θ)	Gamma$(\alpha,\ \lambda)$	Gamma$(\alpha + n\bar{X},\ \lambda + n)$
Binomial(k, θ)	Beta$(\alpha,\ \beta)$	Beta$(\alpha + n\bar{X},\ \beta + n(k - \bar{X}))$
Normal(θ, σ)	Normal(μ, τ)	Normal$\left(\frac{n\bar{X}/\sigma^2 + \mu/\tau^2}{n/\sigma^2 + 1/\tau^2}, \frac{1}{\sqrt{n/\sigma^2 + 1/\tau^2}}\right)$

Bayesian Estimate

$$\hat{\theta}_{\mathrm{B}} = \mathbf{E}\{\theta|\boldsymbol{X} = \boldsymbol{x}\} = \begin{cases} \sum_{\theta} \theta\pi(\theta|\boldsymbol{x}) \\ \int_{\theta} \theta\pi(\theta|\boldsymbol{x})d\theta \end{cases}$$

depending on discrete or continuous posterior

The variance gives posterior risk

$$\rho(\hat{\theta}) = \mathrm{Var}\{\theta|\boldsymbol{x}\}$$

Bayesian Credible set

DEFINITION 10.5

Set C is a $(1-\alpha)100\%$ **credible set** for the parameter θ if the posterior probability for θ to belong to C equals $(1-\alpha)$. That is,

$$\boldsymbol{P}\{\theta \in C \mid \boldsymbol{X}=\boldsymbol{x}\} = \int_C \pi(\theta|\boldsymbol{x})d\theta = 1-\alpha.$$

If the posterior $\pi(\theta|x)$ is Normal (or can be approximated as Normal), This is given by.

$$\mu_x \pm z_{\alpha/2}\tau_x = [\mu_x - z_{\alpha/2}\tau_x, \mu_x + z_{\alpha/2}\tau_x]$$

Bayesian Inference

- Calculate Posterior distribution $\pi(\theta|x)$
- Identify H_0 and H_A
- If $P\{H_0\}$ is greater than $P\{H_A\}$ according to $\pi(\theta|x)$ then accept H_0. Else, reject H_0.

Simulating sampling of Random Variable on the basis of samples from U(0,1)

- ♦ Basic distributions:
  - Bernoulli(p)
    1) If u < p return 1 else return 0
  - Binomial(n, p)
    1) Generate n samples from Bernoulli(p)
    2) Count the number of '1' samples
  - Geometric(p)
    1) Keep generating samples from Bernoulli(p) till '1' sample is generated
    2) Return number of samples generated
  - Negative-Binomial(k, p)
    1) Generate k samples from Geometric(p)
    2) Add the values together

- ♦ Discrete distributions
  - Method 1
    1) Generate u
    2) Find i such that F(i-1)<= u < F(i), where F(x) is the cumulative distribution function
  - Method 2
    1) Generate u
    2) Find the smallest possible value of i such that F(i) > u, where F(x) is the cumulative distribution function

- Continuous distributions
  - Method 1 (Rejection Method)
    1) Find a, b, x such that a,b and 0,c forms a bounding box on f(x) where f(x) is the probability distribution function [for all a<= x <=b, 0 <= f(x) <= c]
    2) Generate u_1, u_2
    3) X = a + (b-a) u_1 and Y = cu_2
    4) If Y <= f(X) accept X as the desired sample. Else return to step 2
  - Method 2 (Inverse Transform Method)
    1) Generate u
    2) Return $F^{-1}(u)$ where $F^{-1}(x)$ is the inverse of F(x), the cumulative density function

- Inverse Transform Methods
  - Geometric: $X = \left\lceil \ln(1-U) / \ln(1-p) \right\rceil$
  - Exponential: $X = -1/\lambda \ln(1-U)$
  - Gamma: Generate α samples from Exponential(λ) and add them

- Special Methods
  - Uniform(a, b)
    1) Generate u
    2) Return u * (b-a) + a
  - Poisson (λ)
    1) Generate u_1, u_2 …
    2) Find the largest value k for which $u_1 * u_2 * \ldots * u_k >= e^{-\lambda}$
    3) Return k
  - Normal(μ, σ) [Box-Mueller Transform]
    1) Generate u_1, u_2
    2) $z_1 = \sqrt{-2\ln(u_1)}\cos(2\pi u_2)$
    3) $z_2 = \sqrt{-2\ln(u_1)}\sin(2\pi u_2)$
    4) $x_1 = z_1 \sigma + \mu$
    5) $x_2 = z_2 \sigma + \mu$

Monte Carlo Methods

Represent any complex distribution in terms of simpler distributions and use the given methods to generate samples