1. (25 points) Consider the linear model

$$Y = X\beta + Z\gamma + \epsilon,$$

where $E(\epsilon) = 0$ and $\mathrm{Cov}(\epsilon) = V$. V is assumed known and positive definite, and (β, γ) are unknown. Further, let $A = X(X'V^{-1}X)^{-}X'V^{-1}$, where $-$ denotes generalized inverse, X is $n \times p$, Z is $n \times q$, and both X and Z may be less than full rank. Let $C(H)$ denote the usual label for the column space of an arbitrary matrix H.

(a) (2 points) Show that $(I - A)'V^{-1}(I - A) = (I - A)'V^{-1} = V^{-1}(I - A)$.

(b) (3 points) Show that A is the projection operator onto $C(X)$ along $C(V^{-1}X)^{\perp}$.

(c) (4 points) Let B denote the projection operator onto $C(X, Z)$ along $C(V^{-1}(X, Z))^{\perp}$. Assume that all matrix inverses exist. Show that

$$B = A + (I - A)Z\left[Z'(I - A)'V^{-1}(I - A)Z\right]^{-1} Z'(I - A)'V^{-1}.$$

(d) (5 points) Show that $(\hat{\gamma}, \hat{\beta})$ are generalized BLUE's for the linear model, where $(\hat{\gamma}, \hat{\beta})$ satisfy

$$\hat{\gamma} = \left[Z'(I - A)'V^{-1}(I - A)Z\right]^{-1} Z'(I - A)'V^{-1}(I - A)Y,$$

and

$X\hat{\xi} = AY$

$$X\hat{\beta} = A(Y - Z\hat{\gamma}).$$

A is the PO onto X?

don't need to orthogonalize here since $\hat{\beta}$ in terms of $\hat{\gamma}$

(e) (5 points) Suppose that $\epsilon \sim N_n(0, V)$ and V is known. Further, suppose that (β, γ) are both estimable. From first principles, derive the likelihood ratio test for the hypothesis $H_0 : \gamma = 0$, where (β, γ) are both unknown, and state the exact distribution of the test statistic under the null and alternative hypotheses.

(f) (6 points) Suppose that $\epsilon \sim N_n(0, \sigma^2 R)$, where R is known and positive definite, and $(\beta, \gamma, \sigma^2)$ are all unknown. Further, assume that (β, γ) are both estimable. Derive an exact joint 95% confidence region for $(\beta, \gamma, \sigma^2)$.

1. Consider the linear model $Y = X\beta + Z\gamma + \varepsilon$

where $E(\varepsilon)=0$, $Cov(\varepsilon)=V$ known & pos. def. (β, γ) unknown

Let $A = X(X'V^{-1}X)^{-}X'V^{-1}$ where $-$ is a gen. inverse $X_{n\times p}$, $Z_{n\times q}$ $r(X)\le p$ $r(Z)\le q$

(a) Show $(I-A)'V^{-1}(I-A) = (I-A)'V^{-1} = V^{-1}(I-A)$

① Show A is a projection operator: $A = A^2$

$A^2 = A\cdot A = X\underbrace{(X'V^{-1}X)^{-}X'V^{-1}X(X'V^{-1}X)^{-}}_{=(X'V^{-1}X)^{-} \text{ by properties of g-inv.}}X'V^{-1}$

$= X(X'V^{-1}X)^{-}X'V^{-1} = A$ ✓

② Show $(I-A)$ is a projection operator $(I-A) = (I-A)^2$

$(I-A)^2 = (I-A)(I-A) = I^2 - A - A + A^2 = I - A - A + A = I - A$ ✓

③ $(I-A)'V^{-1}(I-A) = (I-A)'V^{-1}$ mult. both on ~~left~~ right by $(I-A)$

$(I-A)'V^{-1}(I-A)(I-A) = (I-A)'V^{-1}(I-A)$

$(I-A)'V^{-1}(I-A) = (I-A)'V^{-1}(I-A)$ ✓

Thus $(I-A)'V^{-1}(I-A) = (I-A)'V^{-1}$

④ $(I-A)'V^{-1}(I-A) = V^{-1}(I-A)$ mult. both on left by $(I-A)'$

$(I-A)'(I-A)'V^{-1}(I-A) = (I-A)'V^{-1}(I-A)$

$(I-A)'V^{-1}(I-A) = (I-A)'V^{-1}(I-A)$ ✓

Thus $(I-A)'V^{-1}(I-A) = V^{-1}(I-A)$

$(I-A)'(I-A)'$
$= ((I-A)(I-A))'$
$= (I-A)'$

$\therefore (I-A)'V^{-1}(I-A) = (I-A)'V^{-1} = V^{-1}(I-A)$

1. (b) Show that A is the projection operator onto $C(X)$ along $C(V^{-1}X)^{\perp}$. | 2017 Theory 2

In (a) we showed A is a projection operator, so we must show the onto & along spaces.

<u>onto</u>: NTS $AX = X$

V is non-singular & can be written as $V = QQ'$, Q - non-singular.

Consider the OPO on $Q^{-1}X$

$$P = (Q^{-1}X)((Q^{-1}X)'(Q^{-1}X))^{-}(Q^{-1}X)'$$
$$= (Q^{-1}X)(X'(QQ)^{-1}X)^{-}X(Q^{-1\prime})$$
$$= (Q^{-1}X(X'V^{-1}X)^{-}XQ^{-1})$$

By def. of OPOs: $P(Q^{-1}X) = Q^{-1}X$

$$Q^{-1}X(X'V^{-1}X)^{-}XQ^{-1}Q^{-1}X = Q^{-1}X$$
$$Q^{-1}X(X'V^{-1}X)^{-}X(QQ')^{-1}X = Q^{-1}X$$
$$Q^{-1}X(X'V^{-1}X)^{-}XV^{-1}X = Q^{-1}X$$
$$Q^{-1}AX = Q^{-1}X$$
$$AX = X \checkmark$$

$\therefore$ A is a projection onto $C(X)$

<u>along</u> WTS $Aw = 0$ where $w \in C(V^{-1}X)^{\perp}$

let $w \in C(V^{-1}X)^{\perp}$, then $X'V^{-1}w = 0$ by definition

$$Aw = X(X'V^{-1}X)^{-}\underbrace{X'V^{-1}w} = X(X'V^{-1}X)^{-}0 = 0$$

$\therefore$ A projects along $C(V^{-1}X)^{\perp}$.

1.(c) Let B denote the projection operator onto $C(X,Z)$ along $C(V^{-1}(X,Z))^{\perp}$. Assume all matrix inverses exist. Show that

$$B = A + (I-A)Z[Z'(I-A)V^{-1}(I-A)Z]^{-1}Z'(I-A)'V^{-1}$$

we will take B & show it is a ① projection operator ② onto & ③ along

① $B = B^2$

$$B^2 = (A + (I-A)Z[Z'(I-A)V^{-1}(I-A)Z]^{-1}Z'(I-A)'V^{-1})(A + (I-A)Z[Z'(I-A)V^{-1}(I-A)Z]^{-1}Z'(I-A)'V^{-1})$$

$$= A^2 + 2A(I-A)Z[Z'(I-A)V^{-1}(I-A)Z]^{-1}Z'(I-A)'V^{-1}$$

$\boxed{A(I-A)=0}$

$$+ (I-A)Z\cancel{[Z'(I-A)V^{-1}(I-A)Z]^{-1}}\cancel{(Z'(I-A)V^{-1}(I-A)Z)}[Z'(I-A)V^{-1}(I-A)Z]^{-1}(I-A)V^{-1}$$

$$= A + 0 + (I-A)Z(Z'(I-A)V^{-1}(I-A)Z]^{-1}(I-A)V^{-1}$$

$= B$ ✓ $\therefore$ B is a projection operator

② Onto: $B(X,Z) = (X,Z)$

$B(X,Z) = (BX, BZ)$

$$BX = AX + (I-A)Z[Z'(I-A)V^{-1}(I-A)Z]^{-1}Z'(I-A)'V^{-1}X$$

$$= X + (I-A)Z[Z'(I-A)V^{-1}(I-A)Z]^{-1}Z'(I-A)'V^{-1}(I-A)X$$

HOW? $(I-A)'V^{-1}(I-A) = (I-A)'V^{-1}$

$= X + 0 \implies BX = X$

$$BZ = AZ + (I-A)Z[Z'(I-A)V^{-1}(I-A)Z]^{-1}Z'(I-A)V^{-1}Z$$

$$= AZ + (I-A)Z[Z'(I-A)V^{-1}(I-A)Z]^{-1}\underline{Z'(I-A)V^{-1}(I-A)Z}$$

$(I-A)V^{-1} = (I-A)V^{-1}(I-A)$

$= AZ + (I-A)Z$

$= AZ + Z - AZ$

$= Z \implies BZ = Z$ $\therefore$ B is a projection onto $C(X,Z)$

③ Along: $C(V^{-1}(X,Z))^{\perp} = C(V^{-1}X, V^{-1}Z)^{\perp}$ wts $\forall\, w \in C(V^{-1}X, V^{-1}Z)^{\perp}$ $Bw = 0$

we know if $w \in C(V^{-1}X, V^{-1}Z)^{\perp}$ $(X'V^{-1}w, Z'V^{-1}w) = 0$

$$Bw = Aw + (I-A)Z[Z'(I-A)V^{-1}(I-A)Z]^{-1}Z'(I-A)\underbrace{V^{-1}Zw}_{0}$$

$$= Aw = X'(X'V^{-1}X)^{-1}\underbrace{X'V^{-1}w}_{0} = 0$$

$\therefore$ B is a projection operator onto $C(X,Z)$ along $C(V^{-1}(X,Z))^{\perp}$.

1.(d) Show that $(\hat{\gamma}, \hat{\beta})$ are generalized BLUEs for the linear model where $(\hat{\gamma}, \hat{\beta})$ satisfy

$\hat{\gamma} = [Z'(I-A)'V^{-1}(I-A)Z]^{-1} Z'(I-A)'V^{-1}(I-A)Y$

and $X\hat{\beta} = A(Y - Z\hat{\gamma})$

$X\hat{\beta} = MY$ where M is OPO

2017 Theory 2

$(X, Z) = D \quad \xi = (\beta, \gamma)$

$\hat{\xi} = (D'D)^{-1} D'Y$

$D'D\hat{\xi} = D'Y$

Write $V = QQ'$, Q nonsingular.

$\therefore$ the model $Y = X\beta + Z\gamma + \varepsilon$ can be written as

$Q^{-1}Y = Q^{-1}X\beta + Q^{-1}Z\gamma + Q^{-1}\varepsilon$ let this model be known as

$Y^* = X^*\beta + Z^*\gamma + \varepsilon^*$ where $E[\varepsilon^*] = 0$ $var(\varepsilon^*) = I$.

$\therefore$ By Gauss-Markov, the BLUEs can be found w/ the normal equations By GM, use LSE $\hat{\beta} = (X'X)^{-1}X'Y$ $\quad X'X\hat{\beta} = X'Y$

$$\begin{pmatrix} X^{*\prime} \\ Z^{*\prime} \end{pmatrix} (X^* \; Z^*) \begin{pmatrix} \beta \\ \gamma \end{pmatrix} = \begin{pmatrix} X^{*\prime} \\ Z^{*\prime} \end{pmatrix} Y^* \equiv \begin{pmatrix} X^{*\prime}X^* & X^{*\prime}Z^* \\ Z^{*\prime}X^* & Z^{*\prime}Z^* \end{pmatrix} \begin{pmatrix} \beta \\ \gamma \end{pmatrix} = \begin{pmatrix} X^{*\prime} \\ Z^{*\prime} \end{pmatrix} Y^*$$

① $X^{*\prime}X^*\beta + X^{*\prime}Z^*\gamma = X^{*\prime}Y^* \equiv (X'V^{-1}X)\beta + (X'V^{-1}Z)\gamma = X'V^{-1}Y$

② $Z^{*\prime}X^*\beta + Z^{*\prime}Z^*\gamma = Z^{*\prime}Y^* \equiv (Z'V^{-1}X)\beta + (Z'V^{-1}Z)\gamma = Z'V^{-1}Y$

From ① $(X'V^{-1}X)\beta = X'V^{-1}Y - (X'V^{-1}Z)\gamma$

$(X'V^{-1})(X\beta) = X'V^{-1}Y - (X'V^{-1})Z\gamma$

$X(X'V^{-1}X)^{-}X'V^{-1}(X\beta) = X(X'V^{-1}X)^{-}X'V^{-1}Y - X(X'V^{-1}X)^{-}X'V^{-1}Z\gamma$

$AX\beta = AY - AZ\gamma$

$\boxed{X\hat{\beta} = A(Y - Z\hat{\gamma})}$

From ② $(Z'V^{-1}X)\beta + (Z'V^{-1}Z)\gamma = Z'V^{-1}Y$

$Z'V^{-1}A(Y - Z\hat{\gamma}) + (Z'V^{-1}Z)\hat{\gamma} = Z'V^{-1}Y$

$Z'V^{-1}AY - Z'V^{-1}AZ\hat{\gamma} + Z'V^{-1}Z\gamma = Z'V^{-1}Y$

$(Z'V^{-1}Z - Z'V^{-1}AZ)\hat{\gamma} = (Z'V^{-1} - Z'V^{-1}A)Y$

$(Z'V^{-1}(I-A)Z)\hat{\gamma} = (Z'V^{-1}(I-A))Y$

$Z'(I-A)V^{-1}(I-A)Z\hat{\gamma} = Z'V^{-1}(I-A)Y$

$\Rightarrow \hat{\gamma} = (Z'(I-A)V^{-1}(I-A)Z)^{-1}(Z'V^{-1}(I-A))Y$

$\boxed{\hat{\gamma} = (Z'(I-A)V^{-1}(I-A)Z)^{-1} Z'(I-A)V^{-1}(I-A)Y}$

1.(e) Suppose that $\varepsilon \sim N(0,V)$ and V is known, (β,γ) both estimable. From first principles derive the LRT for the hypothesis $H_0: \gamma = 0$, where (β,γ) are both unknown. State the exact dist. of the test under null & alternative hypotheses. | 2017 Theory 2

$Y = X\beta + Z\gamma + \varepsilon = W\delta + \varepsilon$ where $W = (X, Z)$ $\delta = \binom{\beta}{\gamma}$

Since V is non-singular, $V = QQ'$ where Q is nonsingular

Then $Q^{-1}Y = Q^{-1}W\delta + Q^{-1}\varepsilon$

We will assume X & Z are full rank ~~thus~~ and W is full rank

$Y^* = W^*\delta + \varepsilon^*$ where $\varepsilon^* \sim N_n(0, I)$

$Y^* = Q^{-1}Y$, $W^* = (Q^{-1}X, Q^{-1}Z) = (X^*, Z^*)$

Numerator: null

$\exp[-\frac{1}{2}(Y^* - X^*\hat\beta)'(Y^* - X^*\hat\beta)]$ where $\hat\beta$ satisfies $X^*\hat\beta = M_0^* Y$

$= \exp[-\frac{1}{2}(Y^* - M_0^*Y^*)'(Y^* - M_0^*\hat\beta)]$ where $M_0^* = X^*(X^{*\prime}X^*)^{-1}X^{*\prime}$

$\Rightarrow \hat\beta = (X^{*\prime}X^*)^{-1}X^{*\prime}Y$

$= \exp[-\frac{1}{2}Y^{*\prime}(I - M_0^*)Y^*]$

Denominator: alternative

$\exp[-\frac{1}{2}(Y^* - W^*\hat\delta)'(Y^* - W^*\hat\delta)]$ where $\hat\delta$ satisfies $W^*\hat\delta = M^*Y$

$= \exp[-\frac{1}{2}(Y^* - M^*Y)'(Y^* - M^*Y)]$ $M^* = W^*(W^{*\prime}W^*)^{-1}W^{*\prime}$

$\Rightarrow \hat\delta = (W^{*\prime}W^*)^{-1}W^{*\prime}Y$

$= \exp[-\frac{1}{2}Y^{*\prime}(I - M^*)Y^*]$

$$\therefore LRT = \Lambda = \frac{\exp\{-\frac{1}{2}Y^{*\prime}(I - M_0^*)Y^*\}}{\exp\{-\frac{1}{2}Y^{*\prime}(I - M^*)Y^*\}} = \exp\{-\tfrac{1}{2}Y^{*\prime}(I - M_0^*)Y^*\}\exp\{\tfrac{1}{2}Y^{*\prime}(I - M^*)Y^*\}$$

$\log\Lambda = -\frac{1}{2}Y^{*\prime}(I - M_0^*)Y^* + \frac{1}{2}Y^{*\prime}(I - M)^*Y^*$

$-2\log\Lambda = Y^{*\prime}(I - M_0^*)Y^* - Y^{*\prime}(I - M^*)Y^*$

$= Y^{*\prime}(I - M_0^* - I - M^*)Y^*$

$= Y^{*\prime}(M^* - M_0^*)Y^*$ $r(M^* - M_0^*) \overset{\text{as } M^* \,\&\, M_0^* \text{ are OPOs}}{=} r(M^*) - r(M_0^*) = p + q - q = p$

We reject H_0 when $Y^{*\prime}(M^* - M_0^*)Y^* > C$

under the null, $Y^{*\prime}(M^* - M_0^*)Y^* \sim \chi^2_p$

under the alternative, $Y^{*\prime}(M^* - M_0^*)Y^* \sim \chi^2_{p,\gamma}$

where $\gamma = \frac{\|(I - M_0^*)W^*\delta\|^2}{2}$ is the non-centrality parameter

$\therefore$ let $C = \underbrace{\chi^2_p(1-\alpha)}_{1-\alpha^{th} \text{ percentile of } \chi^2_p}$ $\therefore$ reject H_0 if $Y^{*\prime}(M^* - M_0^*)Y^* > \chi^2_p(1-\alpha)$

1. (f) Suppose $\varepsilon \sim N_n(0, \sigma^2 R)$ where R is known, pos. def. $(\beta, \gamma, \sigma^2)$ all unknown. Assume both (β, γ) are estimable $= r(W) = p+q$

Derive an exact joint CR for $(\beta, \gamma, \sigma^2)$

As previously, we can write $R = QQ'$ where Q is non-singular

$$Q^{-1}Y = Q^{-1}W\delta + Q^{-1}\varepsilon \Rightarrow Y^* = W^*\delta + \varepsilon^* \qquad \varepsilon^* \sim N(0, \sigma^2 I) \qquad r(W^*) = r(W) = p+q$$

$$Y^* = Q^{-1}Y, \quad W^* = (Q^{-1}X, Q^{-1}Z) = (X^*, Z^*)$$

we know $\dfrac{M^*Y^* - W^*\delta}{\sigma} \sim N(0, \sigma^2)$ and $\dfrac{Y^{*\prime}(I - M^*)Y^*}{\sigma^2} \sim \chi^2_{n-(p+q)}$

and $M^*Y^* \perp (I - M^*)Y$

$\therefore$

$$1-\alpha = P\left(\chi^2_{p+q}(1-a) < \frac{\|M^*Y^* - W^*\delta\|^2}{\sigma^2} < \chi^2_{p+q}(a),\ \chi^2_{n-p-q}(1-b) < \frac{\|Y^{*\prime}(I-M^*)Y^*\|^2}{\sigma^2} < \chi^2_{n-p-q}(b)\right)$$

$$= P\left(\chi^2_{p+q}(1-a) < \frac{\|M^*Y - W^*\delta\|^2}{\sigma^2} < \chi^2_{p+q}(a)\right) P\left(\chi^2_{n-p-q}(1-b) < \frac{\|Y^{*\prime}(I-M^*)Y^*\|^2}{\sigma^2} < \chi^2_{n-p-q}(b)\right)$$

where (a, b) s.t. $(1-2a)(1-2b) = 1-\alpha$

$\therefore$ The CR is:

$$\left\{(\beta, \gamma, \sigma^2) : \chi^2_{p+q}(1-a) < \frac{\|M^*Y - W^*\delta\|^2}{\sigma^2} < \chi^2_{p+q}(a),\ \chi^2_{n-p-q}(1-b) < \frac{\|Y^{*\prime}(I-M^*)Y^*\|^2}{\sigma^2} < \chi^2_{n-p-q}(b)\right\}$$

2. (25 points) Suppose that the pair (X, Y) is distributed such that $X \sim \text{normal}(0, 1)$ and $Y \sim \text{Bernoulli}(\theta), 0 < \theta < 1$. For example, X could be the log of the level of a certain biomarker and Y an indicator of some disease. Assume that the value of θ is known and given (e.g. we know that the disease prevalence is 0.001).

   Here we try to answer the following question: Given the above specifications, what is the largest possible correlation between X and Y?

   Notation: Define $\rho = \text{corr}(X, Y)$. Use $\phi(\cdot)$ and $\Phi(\cdot)$ to denote the standard normal pdf and cdf, respectively. Define any new notation you use.

   (a) (1 point) Is $\rho = 1$ possible? Explain.

   (b) (6 points) Find the joint distribution for (X,Y) that maximizes corr(X,Y) subject to the model stated above. Show that it (that distribution) has the property that $E[X|Y = 1] = \theta^{-1}\phi(\Phi^{-1}(1 - \theta))$.

   (c) (6 points) Obtain an explicit expression for $\rho^* = \text{corr}(X, Y)$ within the joint distribution found in the previous part. Compute the numerical value of ρ^* for the case $\theta = 0.001$.

   (d) (6 points) Now suppose we are interested in various diseases with different prevalences (θ) ranging in $(0, 1)$. Find the value of θ that leads to the largest possible value of ρ^*, and compute that largest value, to be denoted ρ^{**} (compute its numerical value).

   (e) (3 points) Note: This part is totally independent of the previous parts, even though the models have some similarity. You can reuse results obtained above if needed.

   Suppose that the iid pairs $(X_i, Y_i), i = 1, \ldots, n$, are distributed such that $X_i \sim \text{normal}(0, 1)$, $Y_i \sim \text{Bernoulli}(\theta), 0 < \theta < 1$, and $\text{corr}(X_i, Y_i) = \rho$, where both θ and ρ are unknown parameters. Develop an estimating equation for (ρ, θ) based on the vectors $Z_i := (T_i, Y_i)^\top, i = 1, \ldots, n$, where $T_i := X_i Y_i$. Obtain the estimates $(\hat{\rho}, \hat{\theta})$ in explicit form.

   (f) (3 points) Based on $\hat{\rho}$ from the previous part, develop a large-sample (as $n \to \infty$) 95% confidence interval for ρ. The interval should not depend on any unknown parameters. Describe and justify your procedure clearly.

Note: $\Phi(-3.09) \approx 0.001$, $\Phi(-2.33) \approx 0.01$, $\Phi(-1.96) \approx 0.025$, $\Phi(-1.64) \approx 0.05$, $\Phi(-1.28) \approx 0.1$

2. Suppose (X,Y) is a pair of RV s.t. $X \sim N(0,1)$ $Y \sim \text{Bernoulli}(\theta)$ | 2017 Section 2

$0<\theta<1$, θ assumed known.

Given the above specifications, what is the largest possible correlation between X and Y?

(a) Is $\rho=1$ possible? Explain.

$$\rho = \text{corr}(X,Y) = \frac{\text{Cov}(X,Y)}{\sqrt{\text{Var}(X)\text{Var}(Y)}}$$

$$\rho = \frac{\text{Cov}(X,Y)}{\sqrt{1\cdot\theta(1-\theta)}}$$

$\rho=1 \Rightarrow X = aY$ This linear relationship is impossible, as X could only take 2 values $\nRightarrow$ Normal

$\Rightarrow X$ is

If you know Y & ρ is known, X can only take 2 values $\Rightarrow X \nsim$ Normal

(b) Find the joint dist for (X,Y) that maximizes $\text{corr}(X,Y)$

The denominator is fixed, so we must maximize $\text{Cov}(X,Y)$

$$\text{Cov}(X,Y) = E[XY] - E[X]E[Y] \qquad E[X], E[Y] \text{ are fixed, } E[X]=0 \ E[Y]=\theta$$

$$= E[XY]$$

$$= E[XY \mid Y=0]P(Y=0) + E[XY \mid Y=1]P(Y=1)$$

$$= (1-\theta)E[0 \mid Y=0] + \theta E[X \mid Y=1]$$

$$= \theta E[X \mid Y=1] \qquad \text{we must maximize } E[X \mid Y=1]$$

$E[X \mid Y=1]$ is maximized by putting all the density at higher values of X when $y=1$ and lower values of X when $y=0$.

for example: $\{(x,y): x<c, y=0\} \cup \{(x,y): x>c, y=1\}$ $\quad Y = I(X>c)$ $P(Y=1) = 1-\Phi(c)$

$$\theta = 1-\Phi(c)$$

$$1-\theta = \Phi(c)$$

$$c = \Phi^{-1}(1-\theta)$$

$$Y = I(X > \Phi^{-1}(1-\theta))$$

$$E[X \mid Y=1] = \frac{E[X \wedge Y=1]}{E[Y=1]} = \frac{\phi(\Phi^{-1}(1-\theta))}{\theta}$$

$$\{(x,y): x<\Phi^{-1}(1-\theta), y=0 \ \cup \ x>\Phi^{-1}(1-\theta), y=1\}$$

$$\begin{cases} \text{when } y=0 \text{ then } X \sim N(0,1)\,I(x<\Phi^{-1}(1-\theta)) \\ \text{when } y=1 \text{ then } X \sim N(0,1)\,I(x\geq\Phi^{-1}(1-\theta)) \end{cases} \text{pdf}$$

2(c). Obtain an explicit expression for $\text{corr}(X,Y) = \rho^*$ w/ the joint distribution found in (b), with $\theta = 0.001$.

$$\rho^* = \frac{\text{Cov}(X,Y)}{\sqrt{\theta(1-\theta)}} = \frac{E[XY]}{\sqrt{\theta(1-\theta)}} = \frac{\theta E[X|Y=1]}{\sqrt{\theta(1-\theta)}} = \frac{\theta \frac{1}{\theta}\phi(\Phi^{-1}(1-\theta))}{\sqrt{\theta(1-\theta)}} = \frac{\phi(\Phi^{-1}(1-\theta))}{\sqrt{\theta(1-\theta)}}$$

when $\theta = 0.001$,

$$\rho^* = \frac{\phi(\Phi^{-1}(.999))}{\sqrt{.001(.999)}} = \frac{\phi(3.09)}{\sqrt{.001(.999)}} = \frac{\frac{1}{\sqrt{2\pi}}\exp\{-\frac{3.09^2}{2}\}}{\sqrt{.001(.999)}}$$

(d) Now suppose we are interested in all possible $\theta \in (0,1)$. Find the value of θ that leads to the largest possible value of ρ^* & compute ρ^* @ that value.

we want to maximize $\phi(\Phi^{-1}(1-\theta))$ and minimize $\sqrt{\theta(1-\theta)}$

The maximum for $\phi(\Phi^{-1}(1-\theta))$ occurs when $\Phi^{-1}(1-\theta) = 0 \Rightarrow \theta = .5$

we can see this also minimizes $\sqrt{\theta(1-\theta)}$ a $\theta - \theta^2$

$$\frac{\partial(\theta-\theta^2)}{\partial\theta} = 1 - 2\theta \overset{set}{=} 0 \Rightarrow \theta = 1/2 \qquad \frac{\partial^2(\theta-\theta^2)}{\partial\theta^2} = -2 < 0 \therefore \text{minimum @ } 0.5$$

$\therefore$ 0.5 simultaneously maximizes $\phi(\Phi^{-1}(1-\theta))$ & minimizes $\sqrt{\theta(1-\theta)}$

$\therefore \rho^*$ when $\theta = 0.5 \Rightarrow$

$$\rho^* = \frac{\phi(\Phi^{-1}(1-.5))}{\sqrt{.5(1-.5)}} = \frac{\phi(\Phi^{-1}(.5))}{.5} = \frac{\phi(0)}{.5} = \frac{1/\sqrt{2\pi}}{.5}$$

$$= 2/\sqrt{2\pi} = \sqrt{2/\pi}$$

2 (e). Suppose (X_i, Y_i) iid pairs $i=1,\dots,n$ s.t. | 2017 Theory 2

$X_i \sim N(0,1)$, $Y_i \sim \text{Bernoulli}(\theta)$ $0<\theta<1$ $\text{corr}(X_i,Y_i)=\rho$ where both θ, ρ are unknown parameters.

Develop an estimating equation for (ρ, θ) based on the vectors $Z_i := (T_i, Y_i)^T$ $i=1,\dots,n$ where $T_i = X_i Y_i$. Obtain the estimates $(\hat\rho, \hat\theta)$ explicitly.

We know from (b) $\rho = \dfrac{E[XY]}{\sqrt{\theta(1-\theta)}}$

$\therefore$ using means, $\hat\rho = \dfrac{\bar T}{\sqrt{\bar Y(1-\bar Y)}}$ where $\bar T = \frac{1}{n}\sum_{i=1}^{n} T_i$, $\bar Y = \frac{1}{n}\sum_{i=1}^{n} Y_i$

$\hat\theta = \bar Y$

(f). Based on $\hat\rho$ from the previous part, develop a large sample 95% CI for ρ. The interval should not depend on any unknown parameters

$$\text{cov}(XY,Y) = E[XY^2] - E[XY]E[Y] \overset{\text{but } Y^2=Y}{=} E[XY] - E[XY]E[Y]$$
$$= E[XY](1-E[Y]) = E[XY](1-\theta)$$

$\sqrt{n}(\bar T - E(XY)) \xrightarrow{d} N(0, \text{var}(XY))$

$\sqrt{n}(\bar Y - \theta) \xrightarrow{d} N(0, \theta(1-\theta))$ $\rho\sqrt{\theta(1-\theta)}$ as $\rho = \dfrac{E[XY]-E[X]E[Y]}{\sqrt{\text{var}(X)\text{var}(Y)}}$

$$\sqrt{n}\begin{pmatrix}\bar T - E(XY)\\ \bar Y - \theta\end{pmatrix} \xrightarrow{d} N\left(\begin{pmatrix}0\\0\end{pmatrix}, \begin{pmatrix}\text{var}(XY) & \text{cov}(XY,Y)\\ & \theta(1-\theta)\end{pmatrix}\right)$$

let $g(A,B) = \dfrac{A}{\sqrt{B(1-B)}}$ $\nabla g(A,B) = \left(\dfrac{1}{\sqrt{B(1-B)}}, \dfrac{-A(1-2B)}{2(B(1-B))^{3/2}}\right)$

$\frac{\partial}{\partial B} A(B(1-B))^{-1/2} = A(-1/2)(B(1-B))^{-3/2}(1-2B)$

$\therefore \sqrt{n}\left(\dfrac{\bar T}{\sqrt{\bar Y(1-\bar Y)}} - \rho\right) \xrightarrow{d} N(0, \Sigma)$

2(f) con't. $\Sigma = \nabla g(E(XY),\theta)\begin{pmatrix} var(XY) & Cov(XY,Y) \\ & \theta(1-\theta)\end{pmatrix}\nabla g'(E(XY),\theta)$

$$= \left(\frac{1}{\sqrt{\theta(1-\theta)}}, \ -\frac{E(XY)(1-2\theta)}{2(\theta(1-\theta))^{3/2}}\right)\begin{pmatrix} var(XY) & Cov(XY,Y) \\ Cov(XY,Y) & \theta(1-\theta)\end{pmatrix}\begin{pmatrix}\frac{1}{\sqrt{\theta(1-\theta)}} \\ -\frac{E(XY)(1-2\theta)}{2(\theta(1-\theta))^{3/2}}\end{pmatrix}$$

$$\left(\frac{var(XY)}{\sqrt{\theta(1-\theta)}} - \frac{Cov(XY,Y)E(XY)(1-2\theta)}{2(\theta(1-\theta))^{3/2}}, \ \frac{Cov(XY,Y)}{\sqrt{\theta(1-\theta)}} - \frac{E(XY)(1-2\theta)(\theta(1-\theta))}{2(\theta(1-\theta))^{3/2}}\right)\begin{pmatrix}\frac{1}{\sqrt{\theta(1-\theta)}} \\ -\frac{E(XY)(1-2\theta)}{2(\theta(1-\theta))^{3/2}}\end{pmatrix}$$

$$\frac{var(XY)}{\theta(1-\theta)} - \frac{Cov(XY,Y)E(XY)(1-2\theta)}{2(\theta(1-\theta))^2} - \frac{Cov(XY,Y)E(XY)(1-2\theta)}{2(\theta(1-\theta))^2} + \frac{(E(XY)(1-2\theta)(\theta(1-\theta)))^2}{4(\theta(1-\theta))^3}$$

$$= \frac{var(XY)}{\theta(1-\theta)} - \frac{Cov(XY,Y)E(XY)(1-2\theta)}{(\theta(1-\theta))^2} + \frac{[E(XY)(1-2\theta)(\theta(1-\theta))]^2}{4(\theta(1-\theta))^3}$$

using sample cov matrix to estimate $\begin{pmatrix} var(XY) & cov(XY,Y) \\ & \theta(1-\theta)\end{pmatrix}$ & use $\bar{T}$ & $\bar{Y}$ to estimate $E(XY)$, θ.

$$\Rightarrow \begin{pmatrix} \frac{1}{n}\Sigma(T_i-\bar{T})^2 & \frac{1}{n}\Sigma(T_i-\bar{T})(Y_i-\bar{Y}) \\ & \frac{1}{n}\Sigma(Y_i-\bar{Y})^2\end{pmatrix} := \begin{pmatrix}\sigma_T^2 & \sigma_{TY} \\ & \sigma_Y^2\end{pmatrix}$$

$$\Rightarrow \hat{\sigma}^2 = \frac{\sigma_T^2}{\sigma_Y^2} - \frac{\sigma_{TY}\bar{T}(1-2\bar{Y})}{(\sigma_Y^2)^2} + \frac{(\bar{T}(1-2\bar{Y})\sigma_Y^2)^2}{4(\sigma_Y^2)^3}$$

$$\frac{\sqrt{n}}{\sigma}\left(\frac{\bar{T}}{\sqrt{\bar{Y}(1-\bar{Y})}} - p\right) \xrightarrow{d} N(0,1)$$

$\therefore$ a 95% CI for p is:

$$\left\{p : Z_{.025} < \frac{\sqrt{n}}{\sigma}\left(\frac{\bar{T}}{\sqrt{\bar{Y}(1-\bar{Y})}} - p\right) < Z_{.975}\right\} = \left\{p : \frac{\sigma Z_{.025}}{\sqrt{n}} < \frac{\bar{T}}{\sqrt{\bar{Y}(1-\bar{Y})}} - p < \frac{\sigma Z_{.025}}{\sqrt{n}}\right\}$$

$$= \left\{p : \frac{\bar{T}}{\sqrt{\bar{Y}(1-\bar{Y})}} - \frac{\sigma Z_{.975}}{\sqrt{n}} < p < \frac{\bar{T}}{\sqrt{\bar{Y}(1-\bar{Y})}} - \frac{\sigma Z_{.025}}{\sqrt{n}}\right\}$$

$$= \left\{p : \frac{\bar{T}}{\sqrt{\bar{Y}(1-\bar{Y})}} - \frac{\sigma Z_{.975}}{\sqrt{n}} < p < \frac{\bar{T}}{\sqrt{\bar{Y}(1-\bar{Y})}} + \frac{\sigma Z_{.975}}{\sqrt{n}}\right\}$$

where Z_a is the a^{th} percentile of the $N(0,1)$.