

STA 532 Midterm, Spring '23

William Tirone

3/9/23

1)

By submitting this exam paper I certify that I have not given or received any help on this exam and the submitted work is my work alone.

2)

Yes, using the law of total expectation and the law of total variance we can accomplish this.

$$\begin{aligned} E(X) &= E(E(X|Z)) \\ &= E(\mu Z) \\ &= \mu E(Z) \\ &= \mu \frac{\kappa}{\kappa} \quad \text{using fact (ii)} \\ &= \mu \end{aligned}$$

$$\begin{aligned} Var(X) &= E(Var(X|Z)) + Var(E(X|Z)) \\ &= E(\mu Z) + Var(\mu Z) \\ &= \mu + \mu^2 Var(Z) \\ &= \mu + \frac{\mu^2}{\kappa} \end{aligned}$$

And (retroactively) comparing to my answer in part 5 using the given moments, this is correct.

3)

$$\begin{aligned}
M_X(t) &= E(e^{tX}) \\
&= E(E(e^{tX}|Z)) \quad (\text{iterated expectation}) \\
&= E(e^{\mu Z(e^t-1)}) \quad (\text{poisson MGF, given}) \\
&= \int_0^\infty e^{\mu z(e^t-1)} \frac{\kappa^\kappa}{\Gamma(\kappa)} z^{\kappa-1} e^{-\kappa z} dz \\
&= \frac{\kappa^\kappa}{\Gamma(\kappa)} \int_0^\infty e^{\mu z(e^t-1)-\kappa z} z^{\kappa-1} dz \\
&= \frac{\kappa^\kappa}{\Gamma(\kappa)} \frac{\Gamma(\kappa)}{[-(\mu(e^t-1)-\kappa)]^\kappa} \quad (\text{using given fact about integral}) \\
&= \left[\frac{\kappa}{\kappa(1-\frac{\mu}{\kappa}(e^t-1))} \right]^\kappa \\
&= [1 - \frac{\mu}{\kappa}(e^t-1)]^{-\kappa}, \quad -\infty < t < \log(1 + \frac{\kappa}{\mu})
\end{aligned}$$

4)

By uniqueness of the MGF, if we can show that the the MGF of the Negative Binomial converges to $e^{\mu(e^t-1)}$, we will have proved the statement. We also know that $\lim_{n \rightarrow \infty} (1 + \frac{x}{n})^n \rightarrow e^x$. Using this fact:

$$\begin{aligned}
\lim_{\kappa_n \rightarrow \infty} M_{X_n} &= \lim_{\kappa_n \rightarrow \infty} \left[\frac{1}{1 - \frac{\mu}{\kappa}(e^t-1)} \right]^{\kappa_n} \quad (\text{given}) \\
&= \frac{1}{e^{-\mu(e^t-1)}} \\
&= e^{\mu(e^t-1)}
\end{aligned}$$

And since the final line is the MGF of $\text{Poisson}(\mu)$, this shows that $X_n \Rightarrow \text{Poisson}(\mu)$

5)

I'm using the "second approach" hints given in the exam. First, I'll use these properties to find the moments of a Poisson r.v.:

if $X \sim POIS(\lambda)$:

$$E(X) = \lambda$$

$$\begin{aligned} E(X(X-1)) &= E(X^2 - X) \\ &= E(X^2) - E(X) = \lambda^2 \quad (\text{using given fact}) \\ E(X^2) &= \lambda^2 + \lambda \end{aligned}$$

$$\begin{aligned} E(X(X-1)(X-2)) &= E(X^3 - 2X^2 - X^2 + 2X) \\ &= E(X^3) - 3E(X^2) + 2E(X) \\ &= E(X^3) - 3(\lambda^2 + \lambda) + 2\lambda = \lambda^3 \quad (\text{using given fact}) \\ E(X^3) &= \lambda^3 + 3\lambda^2 + \lambda \end{aligned}$$

$$\begin{aligned} E(X(X-1)(X-2)(X-3)) &= E((X^3 - 3X^2 + 2X)(X-3)) \\ &= E(X^4 - 6X^3 + 11X^2 - 6X) \\ &= E(X^4) - 6E(X^3) + 11E(X^2) - 6E(X) = \lambda^4 \quad (\text{using given fact}) \\ E(X^4) &= \lambda^4 + 6E(X^3) - 11E(X^2) + 6E(X) \\ &= \lambda^4 + 6\lambda^3 + 18\lambda^2 + 6\lambda - 11\lambda^2 - 11\lambda + 6\lambda \\ &= \lambda^4 + 6\lambda^3 + 7\lambda^2 + \lambda \end{aligned}$$

Now that we have the moments we can check the first and second moments of $X \sim NBIN(\kappa, \mu)$ to see if we're on the right track.

remembering that $Z \sim \text{Gamma}(\kappa, \kappa)$
 $X|Z \sim \text{Pois}(\mu Z)$

$$E(X) = \mu \quad (\text{as we proved earlier})$$

$$\begin{aligned} E(X^2) &= E[E(X^2)|Z] \\ &= E((\mu Z)^2 + \mu Z) \\ &= \mu^2 E(Z^2) + \mu E(Z) \\ &= \mu^2 \frac{\kappa(\kappa+1)}{\kappa^2} + \mu \quad (\text{using gamma fact}) \\ &= \mu_1 + (1 + \tau)\mu^2 \end{aligned}$$

Now we'll find the 4th moment using the pattern from the previous 3:

$$\begin{aligned}
E(X^4) &= E(E(X^4|Z)) \\
&= E[\mu^4 Z^4 + 6\mu^3 Z^3 + 7\mu^2 Z^2 + \mu Z] \quad (\text{from the 4th moment we found above}) \\
&= \mu^4 E(Z^4) + 6\mu^3 E(Z^3) + 7\mu^2 E(Z^2) + \mu E(Z) \\
&= \mu^4 \frac{\kappa(1+\kappa)(\kappa+2)(\kappa+3)}{\kappa^4} + 6\mu^3 \frac{\kappa(\kappa+1)(\kappa+2)}{\kappa^3} + 7\mu^2 \frac{\kappa(\kappa+1)}{\kappa} + \mu \quad (\text{using gamma fact}) \\
&= \mu + 7(1+\tau)\mu^2 + 6(1+\tau)(1+2\tau)\mu^3 + (1+\tau)(1+2\tau)(1+3\tau)\mu^4 \\
&= \boxed{6}\mu_1 - \boxed{11}\mu_2 + \boxed{6}\mu_3 + (1+\tau)(1+2\tau)(1+3\tau)\mu^4
\end{aligned}$$

My strategy to find the values for the empty boxes was to notice that above we need $6\mu_3$ to get the last term of that expression. But that adds $18\mu_2 - 12\mu_1$ so we then need $-11\mu_2$ to get the last term of μ_2 as $7(1+\tau)\mu^2$. To get back the single μ term, we need $6\mu_1$.

6)

Overdispersion occurs for a random variable X when $Var(X) > E(X)$. Using the moments from question 5, we have:

$$\begin{aligned}
E(X) &= \mu \\
Var(X) &= E(X^2) - [E(X)]^2 \\
&= [\mu + (1+\tau)\mu^2] - \mu^2 \\
&= \mu + \frac{\mu^2}{\kappa}
\end{aligned}$$

So the variance can be thought of as the mean plus a shift controlled by τ . $\tau \rightarrow 0$ as $k \rightarrow \infty$, in which case $E(X) = Var(X)$ so there is no overdispersion. This is confirmed by the statement in question 4 that $X_n \Rightarrow Poisson(\lambda)$, which is a distribution with $E(X) = Var(X)$. Then, with smaller values of κ , $Var(X) > E(X)$ since $\mu + \frac{\mu^2}{\kappa} > \mu$.

7)

With the set $S_t = \{x : R = \sum x_i\}$, conditioning on the statistic R truncates the distribution of X . Knowing X fixes R , but the converse is not true. So we have the follow joint conditional PDF:

$$P_X(X|R=r, \mu, \kappa) = \frac{P(R=r|X, \mu, \kappa)P(X|\mu, \kappa)}{P_R(R|\mu, \kappa)} \quad \text{if } x \in S_T$$

and 0 if $X \notin S_t$, noting that above $P(R = r|X, \mu, \kappa) = 1$.

I'm struggling to evaluate the above pdf, but I think if we could prove that $p(X|R = r, \mu, \kappa) = \frac{P(X|\mu, \kappa)}{P_R(R|\mu, \kappa)} = \text{Multinomial}(r, U)$ that would prove the statement. Alternatively, I'm thinking this must be true because there must be some kind of conjugacy here if we're thinking from a Bayesian perspective where U is our prior and $X|U$ is our likelihood function.

8)

First some code to set up the problem:

```
# set up
n = 36
sum_xi = 1164
sum_xi.two = 39508
sum_xi.three = 1406016
sum_xi.four = 52382116

# problem 8 set up
kappa_star = 100
r_obs = sum_xi
x_bar_obs = sum_xi / n

# MAYBE use maximum mu value, not sure
mu_hat = 32.33
kappa = 1/100

# T statistic function
T_stat = function(X) {
  x_bar = mean(X)
  value = sum((X - x_bar)^2) / x_bar
  return(value)
}
```

We need to rewrite the test statistic to calculate it with our summary statistics:

$$\begin{aligned}
T &= \frac{1}{\bar{X}} \sum_i (X_i - \bar{X})^2 \\
&= \frac{1}{\bar{X}} \sum_i \left[X_i^2 - X_i \bar{X} - \bar{X} X_i + (\bar{X})^2 \right] \\
&= \frac{1}{\bar{X}} \left[\sum_i (X_i^2) - n(\bar{X})^2 \right]
\end{aligned}$$

Now plugging in our values we see $t_{obs} = 57.89691$ from the code below.

```
# calculating value of t_obs
t_obs = (1 / x_bar_obs) * (sum_xi.two - n * (x_bar_obs)^2)
t_obs
```

[1] 57.89691

To sample from the conditional distribution, I used the set up given in question 7. We have fixed κ , but we don't know μ . It's not given in the problem that R is a sufficient statistic so I will prove that here:

Let $X_i \sim \text{NegBin}(\kappa, \mu)$

$$\begin{aligned}
f(x_1, \dots, x_n | \kappa, \mu) &= \prod_{i=1}^n \frac{\Gamma(\kappa + x_i)}{\Gamma(\kappa) x_i!} \left(\frac{\mu}{\mu + \kappa} \right)^{x_i} \left(\frac{\kappa}{\mu + \kappa} \right)^{\kappa} \\
&= \underbrace{\frac{1}{\prod x_i!}}_{h(x)} \underbrace{\frac{\prod \Gamma(\kappa + x_i)}{n \Gamma(\kappa)} \left(\frac{\mu}{\mu + \kappa} \right)^{\sum x_i} \left(\frac{\kappa}{\mu + \kappa} \right)^{n \kappa}}_{g(T(X), \mu, \kappa)}
\end{aligned}$$

By the factorization theorem, R is a sufficient statistic. Because of that, our conditional distribution will be the same $\forall \mu > 0$. For 100,000 MC replicates (1M repeats doesn't seem to want to run on the server), the p-value is approximately ≈ 0.148 as calculated in the code chunk below.

```
set.seed(123)
# MC set up
S = 100000
mc_result = c()

for (i in 1:S){

  # first draw U
```

```

U = rdirichlet(1, alpha=rep(kappa_star,36))

# then draw X
X = rmultinom(1, size=sum_xi, U)

# calculate stat, compare, and store results
stat = T_stat(c(X))
test = stat > t_obs
mc_result = c(mc_result, test)
}

print(mean(mc_result))

```

[1] 0.14751

9)

For $0 \leq \alpha \leq 1$ Casella and Berger define a *level* α test as a test with power function $\beta(\theta)$ if $\sup_{P \in P_0} \beta(\theta) \leq \alpha$, which is in contrast to a *size* α test where $\sup_{P \in P_0} \beta(\theta) = \alpha$. They note that this set of level α tests contains the set of size α tests (hopefully I can cite this book here, the instructions said we could use textbooks but can't use others work. This question seems very definition based so I'm leaving it in). Our set up is:

$$H_0 : \kappa = \kappa^*$$

$$\delta(T) = \text{Reject } H_0 \text{ if } p(x_{obs}, \kappa^*) \leq 2.5\%$$

Further, we can state our test as:

$$P_{\kappa^*}(\delta(T) \text{ rejects } H_0) \leq 2.5\%$$

Which is by definition a rule with $\alpha \leq 2.5\%$. We would also have $97.5\% < \beta_{max} < 100\%$ given our choice of α .

10)

I don't believe we can claim this. I've tried both analytically and with simulations (below) and I keep running into a contradiction. Our p-value is $p(T \geq t | \text{the null is true}) = p(T \geq t | \kappa = \kappa^*)$. With our interval rule, it seems like we're conflating two testing ideas, the p-value and the confidence interval. If we calculate a p-value, we're conditioning on the fact that the null

hypothesis is true. However, if we try to construct a confidence interval, we're trying to let k^* vary.

Below, I attempted to simulate this by simulating the p-values similar to the above approach over a range of κ^* , drawing a value for $\gamma \sim U(0, 330)$ with the upper bound chosen to hopefully capture the upper range of κ values I was trying. This illustrates what I think doesn't make sense, that the confidence interval itself is constructed with the null that conditions on κ^* .

```
ks = c(20:30, 300:330)
contained = c()
for (q in seq(1,100)){
  p_values = c()
  for (j in ks){
    S = 100
    mc_result = c()
    for (i in 1:S){

      # first draw U
      U = rdirichlet(1, alpha=rep(j,36))

      # then draw X
      X = rmultinom(1, size=sum_xi, U)

      # calculate stat, compare, and store results
      stat = T_stat(c(X))
      test = stat > t_obs
      mc_result = c(mc_result, test)
    }
    p_values = c(p_values, mean(mc_result))
  }

  # check if in the interval
  g = runif(1,0,330)
  gamma_lower = ks[ks == min(ks[p_values < 0.975])]
  gamma_upper = ks[ks == max(ks[p_values > 0.025])]

  check_between = between(g, gamma_lower, gamma_upper)
  contained = c(contained, check_between)
}

mean(contained)
```

```
[1] 0.98
```