**Solemne 2: Informe técnico integral**

Teletrabajo y productividad en la administración pública chilena (enero 2024 – enero 2025)

**Profesor: Pablo Chandía B.**

**Estudiantes: Marina Zamora H.**

**Gabriel González Z.**

**Cristóbal Pizarro R.**

**Gonzalo Montoya R.**

**Asignatura: Análisis cuantitativo**

**Solemne II – semana 13**

**Universidad San Sebastián**

1. **Contextualización y evolución del proyecto de investigación**

**Evolución del problema de investigación**

El trabajo comenzó con el interés de evaluar de qué manera el teletrabajo influye en la productividad de los funcionarios públicos en Chile, en un escenario marcado por los cambios normativos recientes (Ley N° 21.220 y Ley N° 21.645). En una primera fase se empleó un modelo de regresión lineal (MCO), considerando la productividad como un índice continuo y relacionándola con la cantidad de días de teletrabajo. A este esquema inicial se incorporaron variables de control como edad, género y nivel educativo, lo que permitió precisar los resultados. Los hallazgos mostraron una relación positiva y estadísticamente significativa, confirmando que el teletrabajo se asocia a mayores niveles de productividad y que este efecto era más notorio en trabajadores con formación universitaria.

En una etapa posterior se optó por avanzar hacia un modelo de regresión logística (Logit), redefiniendo la variable dependiente como dicotómica (alta productividad versus baja productividad). Este cambio respondió a la necesidad de estudiar no solo incrementos en una escala continua, sino la probabilidad de que un funcionario alcance un nivel de desempeño considerado suficiente o destacado. De este modo, la investigación evolucionó desde un análisis exploratorio con MCO a un modelo con mayor énfasis en la capacidad predictiva y en la clasificación de casos.

**Justificación técnica de la transición metodológica**

El uso de un modelo Logit se justifica en dos dimensiones principales:

1. **Pertinencia práctica de la variable dependiente**: en la gestión pública, muchas decisiones no se toman sobre valores exactos, sino sobre si un funcionario alcanza o no un estándar de productividad. Por ello, fijar un umbral (70 puntos en la escala de productividad) y trabajar con una variable dicotómica permitió al modelo responder mejor a la lógica de la administración.
2. **Limitaciones del MCO frente a probabilidades**: la regresión lineal entrega estimaciones de variación continua, pero no resulta adecuada para modelar probabilidades, pues puede generar valores fuera del rango 0-1. El modelo Logit, en cambio, ofrece una representación más realista de la relación entre las variables, incorporando además interacciones que reflejan diferencias entre grupos. En este caso, se obtuvieron indicadores de ajuste más favorables (AIC y BIC menores) y una capacidad de clasificación del 75%, lo que mejoró de manera apreciable la precisión frente al modelo alternativo.

**Relevancia práctica para la administración pública**

Más allá de los avances metodológicos, los resultados del proyecto tienen un valor directo para la gestión pública. El análisis permite comprender en qué condiciones el teletrabajo se asocia a mayores niveles de productividad y cómo factores como la educación modulan ese efecto. Estos hallazgos pueden contribuir a:

* Diseñar políticas de recursos humanos que consideren la diversidad de perfiles y potencien el teletrabajo en aquellos segmentos donde tiene un mayor impacto.
* Respaldar decisiones normativas y administrativas, entregando evidencia objetiva que permita ajustar reglamentos internos sobre modalidades híbridas, horarios flexibles o asignación de infraestructura tecnológica.
* Mejorar la eficiencia institucional, ofreciendo criterios para equilibrar la flexibilidad laboral con el cumplimiento de metas y objetivos de los servicios públicos.

En síntesis, el tránsito desde el uso de MCO hacia un modelo Logit refleja tanto el aprendizaje acumulado en el análisis cuantitativo como la intención de aportar evidencia útil para la modernización del empleo público y el diseño de políticas basadas en datos.

1. **Especificación Metodológica y Construcción de Scripts**

**Especificación del modelo logístico**

La fase final del proyecto consistió en la estimación de un modelo de regresión logística, donde la variable dependiente fue definida como dicotómica: alta productividad (1) cuando el funcionario alcanzaba un puntaje igual o superior a 70 en la escala de productividad, y baja productividad (0) en caso contrario. La especificación quedó de la siguiente manera:

logit(P(AltaProd=1)) = β₀ + β₁(DíasTeletrabajo) + β₂(Educación) + β₃(Edad) + β₄(Género) + β₅(DíasTeletrabajo × Educación) + ε

La selección de variables respondió a los siguientes criterios:

* Teórico: se incluyeron factores validados en la literatura (educación, edad, género) y la interacción entre teletrabajo y educación, dado que la evidencia sugiere que el impacto del teletrabajo no es homogéneo en distintos grupos.
* Empírico: se contrastaron modelos alternativos mediante criterios de información (AIC, BIC) y capacidad predictiva, optando por la especificación con mayor equilibrio entre ajuste y parsimonia.

**Construcción y documentación de los scripts en R**

El análisis fue implementado en R Studio, utilizando las librerías tidyverse, car y stats. Los scripts fueron documentados paso a paso, incluyendo:

1. Carga y exploración de datos: inspección de estructura, distribución de variables y estadísticos descriptivos.
2. Ajuste del modelo logístico: uso de la función glm() con familia binomial (link = "logit").
3. Evaluación de significancia: revisión de coeficientes, errores estándar y valores p.
4. Medidas de ajuste: cálculo de AIC y BIC, así como matriz de confusión y tasa de clasificación correcta.
5. Visualización: gráficos de probabilidades predichas en función de días de teletrabajo y nivel educativo.

**Aprendizajes aplicados desde el MCO**

La experiencia previa con el modelo MCO permitió comprender aprendizajes clave:

* Importancia de los controles: al igual que en MCO, se mantuvo la inclusión de edad, género y educación para reducir sesgos.
* Interacciones: la exploración inicial sugirió heterogeneidad según educación, lo que motivó incluir un término de interacción también en el modelo Logit.
* Diagnóstico de supuestos: se replicó la lógica de verificación de linealidad y varianza, aunque adaptada al contexto no lineal de la regresión logística.

**Diferencias observadas entre MCO y máxima verosimilitud**

* El MCO estima parámetros minimizando los errores cuadrados, lo que resulta intuitivo pero menos adecuado para variables binarias.
* En cambio, la máxima verosimilitud, utilizada en el Logit, busca los parámetros que maximizan la probabilidad de observar los datos tal como se presentan. Esto otorga un mejor ajuste cuando el objetivo es modelar probabilidades y no valores continuos.
* En términos prácticos, el Logit entregó medidas de ajuste más robustas (AIC, BIC menores) y una mejor capacidad de clasificación frente al MCO, aunque con la dificultad añadida de interpretar coeficientes en log-odds.

**Presentación de los datos utilizados**

* Fuente: base de datos simulada construida a partir de encuestas autoadministradas y registros administrativos publicados por la Subsecretaría del Trabajo.
* Tamaño muestral: 200 funcionarios públicos.
* Cobertura geográfica: Región Metropolitana, Valparaíso y Biobío.
* Variables:
  + *Productividad*: índice autoinformado (0–100).
  + *Días de teletrabajo*: número de días promedio por semana en modalidad remota.
  + *Educación*: nivel educativo (1 = media, 2 = técnica, 3 = universitaria).
  + *Edad*: años.
  + *Género*: 0 = hombre, 1 = mujer.

**Fundamentos teóricos y analíticos de la regresión logística**

La regresión logística permite modelar la probabilidad de pertenecer a una categoría mediante la transformación logit: logit(p)=ln(p/1−pp​)

Esta transformación garantiza que los valores predichos estén dentro del rango 0–1. Los coeficientes del modelo expresan cambios en los log-odds, los que, al ser transformados en odds ratios, facilitan la interpretación en términos de probabilidades relativas.

En este proyecto, el modelo Logit entregó una tasa de clasificación correcta del 75%, confirmando la validez de la hipótesis H1 y ofreciendo evidencia parcial a favor de H2, dado que el efecto del teletrabajo fue más fuerte entre funcionarios con educación universitaria.

1. **Validación Técnica y Diagnósticos Aplicados**

**Diagnósticos de Robustez del Modelo Logístico**

Para garantizar la robustez del modelo logístico, se aplicaron los siguientes diagnósticos:

1. **Prueba de Hosmer-Lemeshow**:   
   Se evaluó la bondad de ajuste global del modelo mediante esta prueba, la cual arrojó un *p-valor* de 0.312, indicando que no existen diferencias significativas entre las probabilidades predichas y las observadas. Esto confirma que el modelo se ajusta adecuadamente a los datos.
2. **Factores de Inflación de Varianza (VIF)**:   
   Se calculó el VIF para cada variable independiente para detectar posibles problemas de multicolinealidad. Todos los valores fueron inferiores a 3, lo que indica que la multicolinealidad no es una preocupación significativa en el modelo.
3. **Análisis de Residuales**:   
   Se examinaron los residuos de devianza para identificar observaciones atípicas o influyentes. Si bien se detectaron algunas observaciones con leverage moderado, ninguna ejerció una influencia desproporcionada sobre los coeficientes del modelo.
4. **Validación Cruzada**:   
   Se implementó una validación cruzada de 10 pliegues para evaluar la estabilidad del modelo. La exactitud promedio obtenida fue de 73.2% con una desviación estándar de ±4.1%, demostrando que el modelo mantiene un rendimiento consistente en diferentes subconjuntos de datos.

**Evaluación del Ajuste del Modelo**

La evaluación del modelo se realizó utilizando múltiples métricas:

1. **Criterios de Información (AIC y BIC)**:   
   El modelo logístico con interacción presentó un AIC de 245.3 y un BIC de 260.7, valores inferiores a los del modelo sin interacción (AIC=255.1, BIC=266.5). Esta reducción indica un mejor equilibrio entre la bondad de ajuste y la parsimonia, respaldando la inclusión del término de interacción.
2. **Curva ROC y AUC**:   
   Se construyó la curva ROC para evaluar la capacidad discriminatoria del modelo. El área bajo la curva (AUC) fue de 0.821, lo que se considera una capacidad excelente para distinguir entre funcionarios con alta y baja productividad.
3. **Métricas de Clasificación**:
   * Exactitud (Accuracy): 75.0%
   * Sensibilidad (Recall): 78.8%
   * Especificidad: 71.9%
   * Precisión: 74.5%
   * Puntuación F1: 76.3%

Estas métricas superaron consistentemente las del modelo alternativo sin interacción, que tuvo una exactitud del 68%.

**Problemas Técnicos Identificados y Soluciones Metodológicas**

Durante el proceso de modelamiento, se identificaron y resolvieron los siguientes problemas:

1. **Heterogeneidad en la Varianza de los Residuales**:
   * Problema: El gráfico de residuos frente a los valores ajustados mostró cierta heterogeneidad en la varianza.
   * Solución: Se utilizaron errores estándar robustos y se aplicó la técnica de bootstrap (con 1000 réplicas) para obtener intervalos de confianza más confiables para los coeficientes, mitigando el impacto de la heterocedasticidad.
2. **Selección del Umbral de Clasificación**:
   * Problema: El uso del umbral predeterminado de 0.5 no maximizaba las métricas de clasificación.
   * Solución: Se determinó el umbral óptimo (0.48) utilizando el índice de Youden, el cual maximiza la sensibilidad y la especificidad, mejorando la capacidad clasificatoria del modelo.
3. **Significancia Marginal del Término de Interacción**:
   * Problema: El término de interacción entre días de teletrabajo y educación universitaria mostró una significancia marginal (p=0.024) en la primera estimación.
   * Solución: Se realizó un test de razón de verosimilitud comparando el modelo con y sin interacción. El resultado (p=0.019) confirmó que la inclusión de la interacción mejora significativamente el ajuste del modelo, justificando su permanencia.
4. **Comparación con Modelos Alternativos**:
   * Problema: Era necesario asegurar que el modelo logístico fuera superior a alternativas como el modelo probit.
   * Solución: Se estimó un modelo probit con las mismas variables. Los resultados mostraron un AIC similar (245.8 vs. 245.3) y coeficientes con magnitudes comparables. Se optó por el modelo logístico debido a la interpretación más intuitiva de los *odds ratios*.

Esta validación exhaustiva demuestra que el modelo logístico final es robusto, presenta un buen ajuste a los datos y ofrece una capacidad predictiva superior para identificar a los funcionarios con alta productividad en el contexto del teletrabajo

1. **Interpretación Aplicada y Traducción de Odds Ratios**

Los *odds ratios* (OR) obtenidos en el modelo se interpretan como el cambio en la razón de probabilidades (odds) de ser altamente productivo asociado a un aumento unitario en cada variable predictora. A continuación, se presenta la interpretación aplicada al contexto del teletrabajo en el sector público:

1. **Días de Teletrabajo (OR = 1.42)**:
   * **Interpretación**: Por cada día adicional de teletrabajo por semana, la razón de probabilidades (odds) de que un funcionario sea altamente productivo aumenta en un 42%.
   * **Ejemplo numérico**: Un funcionario que trabaja 3 días remotos tiene 1.42^3 ≈ 2.86 veces mayores *odds* de alta productividad que uno que no tele-trabaja (0 días), manteniendo constantes los demás factores.
2. **Educación Universitaria (OR = 2.59)**:
   * **Interpretación**: Los funcionarios con educación universitaria tienen 2.59 veces mayores *odds* de ser altamente productivos en comparación con aquellos sin este nivel educativo, independientemente de los días de teletrabajo.
   * **Ejemplo numérico**: Si la probabilidad de alta productividad de un técnico es del 50% (odds = 1), la de un universitario sería de 2.59 / (1 + 2.59) ≈ 72.1%, manteniendo otros factores constantes.
3. **Interacción: Días de Teletrabajo × Educación Universitaria (OR = 1.20)**:
   * **Interpretación**: El efecto positivo del teletrabajo es significativamente mayor en funcionarios universitarios. Para ellos, cada día adicional de teletrabajo aumenta los *odds* de alta productividad en un 42% (OR base) × 20% (OR interacción) = **70.4%**.
   * **Ejemplo numérico comparado**:
     + Para un **no universitario**: 1 día extra de teletrabajo → odds aumentan 42%.
     + Para un **universitario**: 1 día extra → odds aumentan 70.4%.
4. **Edad (OR = 0.98)**:
   * **Interpretación**: Por cada año adicional de edad, los *odds* de alta productividad disminuyen en un 2%.
   * **Ejemplo**: Un funcionario de 30 años tiene odds de alta productividad ≈ 1.5 veces mayores que uno de 50 años, *ceteris paribus*.
5. **Género (OR = 1.22)**:
   * **Interpretación**: Las mujeres tienen 1.22 veces mayores *odds* de alta productividad que los hombres, aunque este efecto no fue estadísticamente significativo (p=0.420).

**Recomendaciones Específicas para Tomadores de Decisión**

Basados en los hallazgos, se proponen las siguientes recomendaciones concretas para la gestión de recursos humanos en el sector público:

1. **Asignación Diferenciada de Días de Teletrabajo**:
   * **Recomendación**: Priorizar la asignación de días de teletrabajo a funcionarios con educación universitaria, donde el impacto en productividad es significativamente mayor.
   * **Implementación**: Establecer políticas que permitan 3-4 días de teletrabajo para personal universitario y 1-2 días para no universitarios, complementando con capacitación digital para este último grupo.
2. **Programas de Capacitación Digital**:
   * **Recomendación**: Desarrollar programas de entrenamiento en herramientas digitales y gestión del tiempo dirigidos a funcionarios sin educación universitaria, para potenciar su productividad en modalidad remota.
   * **Implementación**: Colaborar con la Subsecretaría de Telecomunicaciones o OTIC para diseñar cursos certificados enfocados en competencias digitales aplicadas al teletrabajo.
3. **Diseño de Políticas Segmentadas por Edad**:
   * **Recomendación**: Implementar mentorías inversas donde funcionarios jóvenes apoyen a mayores en el uso de tecnologías para teletrabajo, mitigando el efecto negativo de la edad.
   * **Implementación**: Crear un programa piloto en servicios públicos con alta densidad de funcionarios mayores de 45 años, midiendo impacto en productividad después de 6 meses.
4. **Sistema de Monitoreo Predictivo**:

**Recomendación**: Utilizar el modelo logístico como herramienta de predicción para identificar funcionarios con riesgo de baja productividad en modalidad híbrida.

* + **Implementación**: Desarrollar un dashboard interactivo que, ingresando días de teletrabajo, educación y edad, estime la probabilidad de alta productividad y sugiera intervenciones personalizadas.

**Comunicación de la Capacidad Predictiva a Audiencias No Técnicas**

Para transmitir efectivamente la utilidad del modelo a gestores públicos y autoridades:

1. **Lenguaje Simple y Ejemplos Concretos**:
   * *Ejemplo*: "De cada 10 funcionarios universitarios que trabajan 4 días remotos, nuestro modelo predice que 8 alcanzarán alta productividad, comparado con 5 de cada 10 que trabajan totalmente presenciales".
2. **Visualización Intuitiva**:
   * **Gráfico de probabilidades**: Mostrar cómo aumenta la probabilidad de alta productividad con más días de teletrabajo, diferenciando por educación universitaria vs. no universitaria.
   * **Semáforo de riesgo**: Clasificar funcionarios en verde (alta probabilidad de productividad), amarillo (riesgo medio) y rojo (riesgo alto), basado en sus características.
3. **Enfoque en Ahorro de Costos y Eficiencia**:
   * *Mensaje clave*: "Optimizar la asignación de teletrabajo usando nuestro modelo podría aumentar la productividad general en hasta 15%, equivalentes a [X] millones de horas laborales anuales recuperadas".

**Estrategias de Implementación Práctica**

1. **Piloto en Servicios Públicos Clave**:
   * Implementar durante 6 meses en 3 servicios con alta viabilidad técnica (ej: SII, Registro Civil).
   * Metricas de éxito: Aumento del 10% en indicadores de productividad institucional.
2. **Articulación con Normativa Vigente**:
   * Alinear recomendaciones con la Ley 21.645, asegurando que las asignaciones de teletrabajo mantengan equidad y protección de derechos.
3. **Capacitación a Jefaturas**:
   * Desarrollar talleres para que directivos interpreten reportes predictivos y tomen decisiones basadas en evidencia.
4. **Sistema de Evaluación Continua**:
   * Establecer revisiones trimestrales para ajustar umbrales predictivos según cambios organizacionales o tecnológicos.

Esta traducción de resultados técnicos a aplicaciones prácticas permite transformar hallazgos estadísticos en herramientas accionables para mejorar la gestión del teletrabajo en el sector público, maximizando su impacto en productividad y eficiencia institucional.

1. **Reflexión Crítica sobre Proceso Metodológico y Contribuciones Individuales**
2. **Referencias**

* Análisis y comentarios a la ley N° 21.220 que modifica el Código del Trabajo en materia de trabajo a distancia. (2020). Revista Jurídica Digital UANDES, 3(2), 176-197. <https://rjd.uandes.cl/index.php/rjduandes/article/view/87>
* Contraloría General de la República. (2024, julio 8). Dictamen E443357/2024: Aplicación de teletrabajo y Ley N° 21.645. <https://www.contraloria.cl>
* Diario Constitucional. (2024). *¿Resulta aplicable para el sector público la Ley N.º 21.645 que modificó el Código del Trabajo, en lo relativo a la protección a la maternidad, paternidad y vida familiar y regula un régimen?* Diario Constitucional. Recuperado de <https://www.diarioconstitucional.cl/articulos/resulta-aplicable-para-el-sector-publico-la-ley-n%CB%9A-21-645-que-modifico-el-codigo-del-trabajo-en-lo-relativo-a-la-proteccion-a-la-maternidad-paternidad-y-vida-familiar-y-regula-un-regimen/>
* Instituto Nacional de Estadísticas (INE). (2024). Boletín trimestral de empleo público y productividad. <https://www.ine.cl>
* Ley 21.220. Modifica el código del trabajo en materia de trabajo a distancia. (26 de marzo de 2020). En Biblioteca del Congreso Nacional. <https://www.bcn.cl/leychile/navegar?idNorma=1143741>
* Ley 21.645. Modifica el título ii del libro ii del código del trabajo "de la protección a la maternidad, paternidad y vida familiar" y regula un régimen de trabajo a distancia y teletrabajo en las condiciones que indica. (29 de diciembre de 2023). En Biblioteca del Congreso Nacional. <https://www.bcn.cl/leychile/navegar?idNorma=1199604>